
Remote oder In-person Usability-Test?

**Methodenvergleich am Beispiel des Metakatalogs
Swissbib**

Bachelorarbeit 2013

Vorgelegt von: Eveline Schmidt

Betreut durch: René Schneider

Eliane Blumer, Forschungsassistentin ACCEPT

Bilingualer Studiengang Information und Dokumentation

Haute école de gestion (HEG) de Genève

Eidesstattliche Erklärung

Diese Bachelorarbeit wird im Rahmen des Abschlussexamens der Haute école de gestion de Genève im Hinblick auf das Erlangen des Diploms „Bachelor of Sciences HES en information documentaire (études bilingues français-allemand)“ erarbeitet. Die Studentin akzeptiert gegebenenfalls die Vertraulichkeitsklausel. Die Verwendung der Schlussfolgerung und Vorschläge, die in der Bachelorarbeit genannt werden (ohne ihrem Wert vorzugreifen), legen weder die Verantwortung des Autors, noch diejenige des Betreuers der Bachelorarbeit, des Experten oder der HEG fest.

„Ich bestätige, dass ich die vorliegende Arbeit alleine verfasst habe, ohne andere Quellen verwendet zu haben als diejenigen, die in der Bibliographie genannt werden.“

Erstellt in Genf, am 19. Juli 2013

Eveline Schmidt

Abstract

Immer neue Formen von Nutzertests werden verwendet, um die Usability von Benutzerschnittstellen zu evaluieren. Die vorliegende Arbeit vergleicht einen *Discount Usability-Test* mit einem *unmoderierten Remote Usability-Test* und analysiert die Benutzerfreundlichkeit des Schweizerischen Metakatalogs Swissbib. Der Vergleich der beiden Usability-Evaluationsmethoden zeigt, inwiefern sich die Einstellungen und das Verhalten der Teilnehmer der beiden Tests unterscheiden und ob es Differenzen beim Identifizieren von Usability-Problemen gibt. Bei der Evaluation von Swissbib werden Nutzerprobleme aufgedeckt und Verbesserungsvorschläge zeigen mögliche Lösungen auf.

Beide Nutzertests waren identisch aufgebaut und wurden mit dem Katalog Swissbib durchgeführt. Sie enthielten sechs Aufgaben von niedrigem Schwierigkeitsgrad, mehrere *Single Ease Questions*, die *System Usability Scale* und weitere Zusatzfragen. Der Discount Usability-Test wurde mit fünf Teilnehmern durchgeführt. Bildschirm, Gesicht und Thinking Aloud Kommentare wurden mit der auf einem Laptop installierten Software *Morae* erfasst. Für den mit *Loop11* durchgeführten Remote Test wurden mithilfe der Diskussionsliste *Swiss-Lib* 101 Personen rekrutiert.

Die Analyse der beiden Nutzertests zeigte, dass die Benutzerfreundlichkeit von Swissbib positiv beurteilt wurde. In verschiedenen Bereichen wie der Datenverarbeitung, der Lesbarkeit, der Architektur, der Terminologie oder dem Zugang zu den Dokumenten könnten aber noch Verbesserungen vorgenommen werden. Es konnten fünf schwerwiegende, zwölf mittelschwere und zehn leichte Probleme identifiziert werden.

Aufgabenzeiten, Anzahl besuchter Seiten und Erfolgsraten unterschieden sich beim Vergleich der beiden Methoden kaum. Teilnehmer des Remote Tests verwendeten aber weniger Aufwand, um Aufgaben ohne klares Ziel zu bearbeiten. Bei der Einschätzung der Schwierigkeit von Aufgaben oder der Benutzerfreundlichkeit konnten ebenfalls keine signifikanten Unterschiede gefunden werden. Teilnehmer des Discount Tests bewerteten jedoch Fragen, die mit den Fähigkeiten des Moderators in Verbindung gebracht werden konnten, besser. Der Discount Test war mit 22 zu 14 gefundenen Nutzungsproblemen effektiver als die Online Methode.

Inhaltsverzeichnis

Eidesstattliche Erklärung	i
Abstract	ii
Inhaltsverzeichnis	iii
Tabellenverzeichnis	v
Abbildungsverzeichnis	v
Einführung	1
1 Ausgangslage	3
1.1 Ziele der Arbeit	3
1.2 Swissbib	3
2 Usability und Usability-Methoden	5
2.1 Usability	5
2.1.1 Definition	5
2.1.2 Usability-Evaluation	6
2.2 Usability-Test	7
2.2.1 In-person Usability-Test	8
2.2.2 Remote Usability-Test	10
2.2.3 Discount Usability-Test vs. unmoderierter Remote Usability-Test	11
3 Forschungsfragen	14
3.1 Subjektive Einstellung der Teilnehmer	14
3.2 Verhalten der Teilnehmer	15
3.3 Effektivität der Methoden	16
4 Durchführung der Nutzertests	17
4.1 Konzeption des Tests	17
4.1.1 Pretest-Fragebogen	17
4.1.2 Aufgaben und Szenarien	17
4.1.3 Folgefragen zu den Aufgaben	18
4.1.4 Posttest-Fragebogen	19
4.2 Umsetzung des Tests – In-person	20
4.2.1 Datenerhebung	20
4.2.2 Vorbereitungen	20
4.2.3 Testablauf	22
4.3 Umsetzung des Tests – Remote	22
4.3.1 Datenerhebung	22
4.3.2 Vorbereitungen	22
4.3.3 Testablauf	24

5	Usability-Probleme Swissbib	25
5.1	Menüleiste	27
5.2	Autovervollständigung.....	27
5.3	Null-Treffermengen	29
5.4	Suchresultate	30
5.5	Trefferliste	32
5.6	Erweiterte Suche	34
5.7	Bestandesinformationen	35
5.8	Anzeige mit Internet Explorer 8	37
5.9	Hilfe-Themen	38
5.10	My swissbib	40
6	Resultate aus dem Methodenvergleich.....	43
6.1	Überprüfung der subjektiven Einstellung	44
6.1.1	Single Ease Question.....	44
6.1.2	System Usability Scale.....	45
6.1.3	Zweiter Posttest-Fragebogen	46
6.1.4	Fazit zur Hypothese 1	47
6.2	Überprüfung des Verhaltens.....	48
6.2.1	Benötigte Zeit	48
6.2.2	Besuchte Seiten	50
6.2.3	Erfolgsrate	51
6.2.4	Fazit zur Hypothese 2	54
6.3	Überprüfung der Effektivität	56
6.3.1	Anzahl und Schweregrad der identifizierten Usability-Probleme.....	59
6.3.2	Beim Remote Test nicht identifizierte Probleme	60
6.3.3	Beim In-person Test nicht identifizierte Probleme.....	61
6.3.4	Fazit zur Hypothese 3	62
7	Fazit	63
	Literaturverzeichnis	65
	Anhang A: Skript für den Testleiter	69
	Anhang B: Pretest.....	71
	Anhang C: Aufgaben und Folgefragen	72
	Anhang D: Posttest-Fragebögen	75
	Anhang E: Einverständniserklärung - In-person Test.....	77
	Anhang F: Rekrutierungs E-Mail - Remote Test	78
	Anhang G: Ausschnitte aus Loop11 - Remote Test	79
	Anhang H: Usability-Probleme	82

Tabellenverzeichnis

Tabelle 1: Unterschiedliche Usability-Evaluationsmethoden	6
Tabelle 2: Teilnehmer des In-person Tests	21
Tabelle 3: Alter der Teilnehmer des Remote Tests	23
Tabelle 4: Geschlecht der Teilnehmer des Remote Tests.....	23
Tabelle 5: Erfahrungen der Teilnehmer des Remote Tests	24
Tabelle 6: Einschätzung des Schwierigkeitsgrads	44
Tabelle 7: Resultate der Posttest-Fragebögen	45
Tabelle 8: Resultat des zweiten Teils des Posttest-Fragebogens	46
Tabelle 9: Durchschnittlich benötigte Zeit für den gesamten Test	48
Tabelle 10: Durchschnittlich benötigte Zeit pro Aufgabe	49
Tabelle 11: Durchschnittliche Anzahl besuchter Seiten.....	50
Tabelle 12: Durchschnittliche Erfolgsraten pro Aufgabe.....	52
Tabelle 13: Erfolgsrate der Aufgabe 4	53
Tabelle 14: Anteil identifizierter Probleme je Schweregrad	59

Abbildungsverzeichnis

Abbildung 1: Single Ease Question	19
Abbildung 2: Beispielfrage der System Usability Scale	19
Abbildung 3: Identifizierte Usability-Probleme nach Anzahl Teilnehmer.....	20
Abbildung 4: Drei Schweregrade für Usability-Probleme	26
Abbildung 5: Menüleiste Swissbib	27
Abbildung 6: Autovervollständigung „nachtzug nach I“	27
Abbildung 7: Fehlende Autovervollständigung „nachtzug nach li“	28
Abbildung 8: Autovervollständigung "du sollst"	28
Abbildung 9: Übernahme der Autovervollständigung „du sollst nicht töten“	28
Abbildung 10: Alternative Suchvorschläge „gran cru“	29
Abbildung 11: Umlautverarbeitung auf Loop11	29
Abbildung 12: Alternative Suchvorschläge „wicki und die graummer“	30
Abbildung 13: Datensatz „Wickie und die Graumänner“ für St. Gallen	30
Abbildung 14: Erste Suchresultate "grand cru"	31

Abbildung 15: Fassettennavigation	31
Abbildung 16: Umschlagbild „Grand Cru“	32
Abbildung 17: Anzeige der Bibliotheken auf der Trefferliste	32
Abbildung 18: Symbol zum Hinzufügen eines Dokuments auf die Merkliste	33
Abbildung 19: Symbol zum Löschen eines Dokuments aus der Merkliste	33
Abbildung 20: Link zur erweiterten Suche	34
Abbildung 21: Auswahl der Formate in der erweiterten Suche	34
Abbildung 22: Optionen für nähere Informationen über ein Dokument.....	35
Abbildung 23: Informationen zum Churer Hörbuch „Wickie und die Graumänner“.	35
Abbildung 24: Button „zum Bestand“	36
Abbildung 25: Buttons „Bibliotheksinformation“ und „Favorit-Bibliothek“	36
Abbildung 26: Anzeige der Verfügbarkeit mit dem Internet Explorer 8	37
Abbildung 27: Anzeige des Marc-Formats mit dem Internet Explorer 8	37
Abbildung 28: Richtige Anzeige des Marc-Formats mit Firefox	37
Abbildung 29: Startseite mit unterschiedlichen Hilfeseiten	38
Abbildung 30: Navigation innerhalb der Hilfe	38
Abbildung 31: Hilfeseite	39
Abbildung 32: Menüoptionen „My swissbib“ und „Anmelden“	40
Abbildung 33: Heatmap "My swissbib" und "Anmelden".....	40
Abbildung 34: Loginvarianten	41
Abbildung 35: Zugriff auf das Benutzerkonto.....	41
Abbildung 36: Favorit-Bibliotheken auf den persönlichen Sucheinstellungen	42
Abbildung 37: Temporäre Merkliste	42
Abbildung 38: Anzeige der Verfügbarkeit - Lokaler Katalog	53
Abbildung 39: Anzeige der Verfügbarkeit - Swissbib	54
Abbildung 40: Problem mit der Funktion „Path-Analysis“	57
Abbildung 41: Fehlermeldung	57
Abbildung 42: Klickstream-Daten aus Loop11	58
Abbildung 43: Heatmap einer spezifischen Seite aus Loop11.....	58
Abbildung 44: Heatmap der Startseite aus Loop11	58
Abbildung 45: Anzahl identifizierte Usability-Probleme.....	59

Einführung

Immer neue Evaluationsmethoden erlauben, die Benutzerfreundlichkeit einer Benutzeroberfläche zu prüfen. Früher war es üblich, solche Tests in Labors durchzuführen. Diese Methode bedurfte intensiver Vorbereitungszeit und passenden Equipments, weshalb Nutzertests noch heute oft als zeitaufwendig und kostenintensiv gelten. Durch den Wunsch nach kostengünstigen und schnellen Evaluationsmethoden werden heute Benutzerfreundlichkeitstests vermehrt in kleinerem Rahmen mit wenigen Teilnehmern durchgeführt. Diese sogenannten *Discount Usability-Tests* sollen den Aufwand reduzieren und die Wahrscheinlichkeit von häufigem Testen erhöhen. Der neueste Trend ist es, Tests über das Internet durchzuführen. Bei diesen *Remote Usability-Tests* können die Teilnehmer den Test durch den Einsatz neuer Technologien zu Hause durchführen. Bei *unmoderierten Remote Usability-Tests* geschieht dies sogar ohne Intervention eines Moderators.

Es stellt sich die Frage, wie diese neueren Methoden die Resultate eines Usability-Tests beeinflussen. Wie gut können sie Usability-Probleme identifizieren und inwiefern wird durch sie das Verhalten und die Einstellung der Teilnehmer während des Tests beeinflusst?

Diese Bachelorarbeit befasst sich mit den zwei Methoden *Discount Usability-Test* und *unmoderierter Remote Usability-Test* und vergleicht die beiden. Dazu wurde derselbe Test einmal persönlich und einmal online durchgeführt. Als Untersuchungsobjekt diente der Metakatalog Swissbib, welcher eine Benutzeroberfläche für die Suche von Bibliothekssammlungen aus der ganzen Schweiz bietet. Drei unterschiedliche Hypothesen über die Unterschiede der beiden Methoden wurden aufgestellt und überprüft. Ein Fokus der Untersuchung lag dabei auf der Effektivität beim Identifizieren von Usability-Problemen. Zudem wurde untersucht, ob die Erfolgsraten oder der investierte Aufwand beim Lösen der Aufgaben variieren. Schliesslich wurde anhand von Fragebögen evaluiert, ob sich die subjektiven Einstellungen je nach Methode unterscheiden.

Neben dem Vergleich der beiden Evaluationsmethoden war die Identifikation von Usability-Problemen bei Swissbib ein weiteres Ziel. Die Probleme, welche bei einem der beiden Tests gefunden wurden, wurden beschrieben und illustriert. Zusätzlich wurden die Nutzungsprobleme nach Schweregrad geordnet und Vorschläge sollen Verbesserungsmöglichkeiten aufzeigen.

Im ersten Kapitel wird auf die Ausgangslage der Arbeit eingegangen. Dazu werden die Ziele sowie die getestete Webseite Swissbib näher beschrieben. Die theoretische Grundlage für die Arbeit befindet sich im zweiten Kapitel. Hier werden Begriffe wie Usability geklärt und die Evaluationsmethoden In-person und Remote Usability-Test beleuchtet. Das darauf folgende Kapitel widmet sich den Forschungsfragen und legt die Hypothesen dar, die der Arbeit zugrunde liegen. Kapitel vier erläutert die praktische Seite der Studie. Es wird erklärt, wie die beiden Nutzertests geplant und wie sie umgesetzt wurden. Unter Punkt fünf und sechs werden die Resultate der Untersuchung besprochen. Zuerst werden die identifizierten Usability-Probleme bei Swissbib dargelegt und Verbesserungsvorschläge gemacht. Die Probleme sind mit Bildern und konkreten Beispielen illustriert. Dann folgen ein Vergleich der beiden Forschungsmethoden und eine Prüfung der Hypothesen. Den Abschluss der Arbeit bildet schliesslich das siebte Kapitel mit einem Fazit.

1 Ausgangslage

1.1 Ziele der Arbeit

Das erste Ziel der Bachelorarbeit besteht darin, die beiden Usability-Methoden *unmoderierter Remote Usability-Test* und *In-person Discount Usability-Test* zu vergleichen. Es soll untersucht werden, inwiefern die Methoden in ihrer Effektivität beim Identifizieren von Usability-Problemen variieren. Ausserdem wird analysiert, ob sich Verhalten oder subjektive Einstellung der Teilnehmer während eines Tests unterscheiden. Dazu wurden drei Hypothesen über mögliche Unterschiede der beiden Methoden verfasst. Zu jeder Hypothese wurden Indikatoren festgelegt, die dazu dienen, die Forschungsfragen zu überprüfen. Ein und derselbe Test wurde einmal persönlich mit einem Moderator und einmal online mit einer speziellen Software durchgeführt. Als Untersuchungsobjekt wurde der Metakatalog Swissbib gewählt, der im nächsten Kapitel näher beschrieben wird. Durch die Analyse der Resultate der beiden Tests sollen die Hypothesen gestützt, falsifiziert oder gegebenenfalls angepasst werden.

Das zweite Ziel der Arbeit besteht darin, Usability-Probleme des Katalogs Swissbib zu identifizieren. Die Nutzungsprobleme sollen grafisch dargestellt, mit Beispielen belegt und durch mögliche Lösungsvorschläge ergänzt werden. Zudem wurde eine Klassifizierung nach Prioritäten erstellt, indem jedem Problem sein Schweregrad zugeordnet wurde.

1.2 Swissbib

Swissbib¹ ist ein gesamtschweizerischer Metakatalog, welcher einen zentralen Einstieg für die Recherche in den Beständen der Schweizer Universitäts- und Hochschulbibliotheken, der Schweizerischen Nationalbibliothek und verschiedener digitaler Repositorien ermöglicht. Das Projekt ist Teil des Innovations- und Kooperationsprojekts E-lib.ch (Elektronische Bibliothek Schweiz), welches zum Ziel hat, einen „einfachen und schnellen Zugang zu Informationsressourcen unterschiedlichster Art“² zu bieten. Das Projekt der Schweizer Hochschul- und Universitätsbibliotheken sowie der Nationalbibliothek wird von der Universitätsbibliothek Basel geleitet. Die heterogenen Datenbestände sollen vereinigt und harmonisiert werden, um auf einer einheitlichen Oberfläche angeboten zu werden. Seit dem Dezember 2009 ist Swissbib als Beta-Version online zugänglich.³

¹ UNIVERSITÄT BASEL. *Swissbib* [online]. 2013. <http://www.swissbib.ch> (letzter Zugriff: 08.06.2013)

² SWISSBIB. E-lib.ch/de. In: *Swissbib project wiki* [online]. <http://www.swissbib.org/wiki/index.php?title=E-lib.ch/de> (letzter Zugriff: 09.06.2013)

³ BÖTTCHER, Uwe. *Swissbib: die ganze Schweiz auf einen Blick. B.I.T* [online]. 2010, Bd. 13, Nr. 2. S. 157-160. <http://www.b-i-t-online.de/heft/2010-02/nachrichtenbeitraege.pdf> (letzter Zugriff: 08.06.2013)

Die technische Infrastruktur von Swissbib wird von OCLC (Online Computer Library Center) bereitgestellt und besteht aus drei unabhängigen Schichten. Das Metadaten-Managementsystem CBS ist für das Sammeln, den Import und die Aufbereitung der Daten verantwortlich. SOLR⁴ ist die Suchmaschine, welche die Daten indexiert und anreichert. Das Informationsportal OCLC TouchPoint wurde für die Realisierung der Benutzerschnittstelle, der sichtbaren Oberfläche für die Nutzer, verwendet. Swissbib führt somit keine simultanen Suchen in den unterschiedlichen Datenquellen durch, sondern sammelt die Daten ein und speichert sie in einer eigenen Datenbank, welche dann für die Suche verwendet wird. Der modulare Aufbau von Swissbib und die Speicherung der Daten in einer eigenen Datenbank ermöglichen es, moderne Web 2.0-Technologien wie Tags oder persönliche Listen zu integrieren.⁵ Benutzer können in 19,4 Millionen Dokumenten unterschiedlichster Sammlungen suchen. Folgende Bibliotheken und Bibliotheksverbünde sind bisher integriert.⁶

- Aargauer Bibliotheksnetz (23 Bibliotheken)
- Bibliotheksverbund Alexandria (20 Bibliotheken)
- Bibliotheksverbund Graubünden (25 Bibliotheken)
- IDS Basel/Bern (197 Bibliotheken)
- IDS Luzern (30 Bibliotheken)
- IDS St. Gallen (37 Bibliotheken)
- NEBIS (97 Bibliotheken)
- Schweizerische Nationalbibliothek (1 Bibliothek)
- Sistema bibliotecario ticinese (31 Bibliotheken)
- St. Galler Bibliotheksnetz (57 Bibliotheken)
- RERO Westschweizer Bibliotheksverbund (236 Bibliotheken)

Ausserdem sind in Swissbib folgende digitale Repositorien eingebunden:

- Catalogue collectif suisse des affiches
- Digitalisate des Eidgenössischen Archivs für Denkmalpflege
- Digitalisate des Schweizerischen Literaturarchives
- retro.seals.ch
- ZORA

Für das Jahr 2013 ist die Integration der Liechtensteinischen Landesbibliothek mit 25 Bibliotheken und der Datenquellen rerodoc, alexandria und serval geplant. Ausserdem sind Neuerungen in den Bereichen Bestellung, Benutzerkonto und Datenaufbereitung in Vorbereitung und weitere Wörterbücher für die Suche sollen eingebunden werden.⁷

⁴ SWISSBIB. Project state/de. In: *Swissbib project wiki* [online].
http://www.swissbib.org/wiki/index.php?title=Project_state/de (letzter Zugriff: 09.06.2013)

⁵ BÖTTCHER, Uwe. *Swissbib: die ganze Schweiz auf einen Blick. B.I.T* [online]. 2010, Bd. 13, Nr. 2. S. 157-160. <http://www.b-i-t-online.de/heft/2010-02/nachrichtenbeitraege.pdf> (letzter Zugriff: 08.06.2013)

⁶ SWISSBIB. Über swissbib. In: *Swissbib* [online].
<http://www.swissbib.ch/TouchPoint/aboutswissbib.do;jsessionid=CED706B2067574E94AC6ED1D974C688C.worker4?methodToCall=showAboutSwissbib&submenue=nav1&lang=de> (letzter Zugriff: 05.04.2013)

⁷ Ebd.

2 Usability und Usability-Methoden

2.1 Usability

2.1.1 Definition

Für Usability werden im Deutschen auch die Begriffe Benutzerfreundlichkeit, Benutzungsfreundlichkeit oder Benutzbarkeit verwendet (Schweibenz, Thissen, 2003: 39). Sie bezeichnet ein Qualitätsmerkmal eines beliebigen technischen Produkts, wobei sich die Qualität auf die Interaktion zwischen einem Benutzer und diesem Produkt bezieht. Usability misst, inwiefern ein Benutzer durch die Nutzung eines Systems seine Bedürfnisse befriedigen kann. Sie ist somit nicht zu verwechseln mit dem Begriff Funktionalität, welcher sich auf die Fähigkeit eines Produkts bezieht, bestimmte Aufgaben zu lösen und unabhängig von der Interaktion mit einem Benutzer ist (Dumas, Redish, 1999: 4).

Für den Begriff Usability existiert eine Vielzahl an Definitionen, welche unterschiedliche Schwerpunkte legen. Die wichtigsten Aspekte sind jedoch immer ein Produkt, ein Benutzer und die Interaktion zwischen diesen beiden (Schweibenz, Thissen, 2003: 41). Dabei ist der Fokus immer auf den Benutzer und dessen Bedürfnisse gerichtet und es ist der Benutzer, der bestimmt, wann ein Produkt die Qualität der Benutzerfreundlichkeit erfüllt oder wann nicht. Denn es sind seine Ziele und Bedürfnisse, die erreicht werden sollen (Dumas, Redish, 1999: 5).

In der internationalen Norm ISO 9241-11 wird Usability definiert: Usability sei das Ausmass, mit welchem ein Produkt in einem spezifischen Nutzungskontext von einem bestimmten Benutzer genutzt werden könne, um dessen Ziele effizient, effektiv und zufriedenstellend zu erreichen. Die Benutzerfreundlichkeit eines Produkts kann nach dieser Definition somit an den drei folgenden Kriterien gemessen werden (Barnum, 2011: 11 f., Schweibenz, Thissen, 2003: 40):

- Effektivität: Nutzer sollen ihre Ziele möglichst genau und vollständig erreichen.
- Effizienz: Das Verhältnis zwischen eingesetztem Aufwand und der Genauigkeit sowie der Vollständigkeit der Zielerreichung sollte stimmen.
- Zufriedenheit: Bei dem Nutzer soll sich ein subjektives Wohlbefinden während der Nutzung des Produkts einstellen.

Unterschiedlichen Usability-Forschern reichen diese drei Dimensionen nicht aus und sie erweitern die ISO-Norm um verschiedene weitere Fassetten. Nielsen (1993: 26) beispielsweise schreibt der Usability fünf Aspekte zu:

- Lernbarkeit: Das Produkt sollte einfach zu erlernen sein.
- Effizienz: Sobald Benutzer den Umgang mit dem Produkt gelernt haben, können sie Aufgaben schnell lösen.
- Einprägsamkeit: Auch nach längerem Nichtgebrauch sollte sich der Nutzer erinnern, wie das Produkt funktioniert.
- Fehler: Das Produkt sollte wenige Fehler enthalten und fehlertolerant⁸ sein.
- Zufriedenheit: Nutzer mögen das Produkt und finden seine Nutzung angenehm.

Usability betrifft jedes technische System, bei dem es zur Interaktion zwischen Mensch und Maschine kommt. Jedoch ist Usability vor allem im Kontext von Software und Benutzerschnittstellen ein Thema und die meisten Definitionen oder Evaluationsmethoden wurden im Rahmen der Softwareentwicklung erarbeitet (Sarodnick, Brau, 2011: 29). Nachfolgend wird der Begriff Usability auch immer in diesem Zusammenhang verwendet und die vorgestellten Methoden beziehen sich auf die Evaluation von Benutzerschnittstellen.

2.1.2 Usability-Evaluation

Usability-Evaluation ist unabdingbar, um zu wissen, ob ein Produkt die Bedürfnisse der Benutzer erfüllt (Schweibenz, Thissen, 2003: 38). Kriterien wie Lernbarkeit oder Effizienz geben den Rahmen vor, um die Usability eines Produkts auf systematische und neutrale Art und Weise zu bewerten (Sarodnick, Brau, 2011: 23). Spezifische Kriterien sollten möglichst früh festgelegt werden, damit die Usability eines Produkts bereits bei seiner Entwicklung berücksichtigt werden kann. Ziele der Usability-Evaluation sind, die Benutzerfreundlichkeit eines fertigen oder halb fertigen Produkts zu bewerten und zu verbessern oder die Entscheidungsfindung zu unterstützen (Nielsen, 1993: 26 f.). Im Laufe der Zeit wurden viele unterschiedliche Methoden entwickelt, welche die Evaluation ermöglichen. Evaluations-Methoden lassen sich in verschiedene Kategorien unterteilen, wobei an dieser Stelle nur vier solcher Kriterien vorgestellt werden, welche für die vorliegende Arbeit relevant sind (siehe Tabelle 1).

Tabelle 1: Unterschiedliche Usability-Evaluationsmethoden

<i>Teilnehmer</i>	<i>Zeitpunkt</i>	<i>Datenqualität</i>	<i>Aufwand</i>
Benutzer	Formativ	Qualitativ	Deluxe
Experten	Summativ	Quantitativ	Discount

⁸ Wenn ein Benutzer einen Fehler begeht, kann er mit wenig Aufwand trotzdem zum gewünschten Ergebnis gelangen.

Ein erster Unterschied besteht darin, wer die Evaluation vornimmt. Expertenorientiert oder analytisch werden jene Methoden genannt, bei welchen Usability-Experten mit ihrem Fachwissen das Produkt bewerten. Hingegen sind bei benutzerorientierten oder empirischen Methoden tatsächliche oder mögliche Benutzer involviert (Schweibenz, Thissen, 2003: 75). Ein weiteres Unterscheidungskriterium ist der Zeitpunkt der Evaluation. Findet diese während der Produktentwicklung statt, handelt es sich um eine formative Methode. Die Bewertung eines fertigen Produkts nennt sich summative Evaluation (Sarodnick, Brau, 2011: 120). Die gewonnenen Daten können quantitativer oder qualitativer Qualität sein. Als quantitativ werden Daten bezeichnet, wenn eine grosse Datenmenge mit statistischen Methoden ausgewertet wird. Qualitative Daten werden aus Evaluationen mit wenigen Teilnehmern gewonnen. Hier werden subjektive Daten gesammelt und interpretiert. Weiter kann je nach betriebenem Aufwand zwischen Discount Usability und Deluxe Usability unterschieden werden (Schweibenz, Thissen, 2003: 81 f.). Als Erfinder der Discount Usability gilt Nielsen (1993, S. 17), der sich mit diesem Schlagwort für die Vereinfachung traditioneller Methoden einsetzt. Diese traditionellen Methoden werden von einigen Autoren als Deluxe Usability bezeichnet, weil sie oftmals mit viel Aufwand geplant und umgesetzt werden. Auch Krug (2006: 134 ff.) befürwortet Usability-Tests mit wenigen Teilnehmern und einfachen Mitteln, da dies die Wahrscheinlichkeit und Häufigkeit der Durchführung von Tests erhöht und Zeit sowie Geld spart.

Es gibt sehr viele unterschiedliche Usability-Evaluationsmethoden, die sich in eine oder auch mehrere Kategorien einordnen lassen. In der vorliegenden Arbeit wird ausschliesslich auf den Usability-Test eingegangen, da diese Methode evaluiert werden soll.

2.2 Usability-Test

Der Usability-Test, auch Nutzertest oder Benutzbarkeitstest genannt, gehört zu den benutzerorientierten Methoden. Er simuliert eine Nutzungssituation, um das individuelle Verhalten eines Benutzers mit einem Produkt zu analysieren. Ein Usability-Test kann sowohl formativ als auch summativ durchgeführt werden. Dumas und Redish (1999: 22) haben fünf Merkmale identifiziert:

1. Das Hauptziel ist es, die Usability eines Produkts zu verbessern.
2. Die Teilnehmer repräsentieren tatsächliche Nutzer.
3. Die Teilnehmer bewältigen authentische Aufgaben.
4. Handlungen und Aussagen der Teilnehmer werden beobachtet und aufgenommen.
5. Daten werden analysiert, echte Probleme werden identifiziert und es werden Änderungsvorschläge gemacht, um diese zu lösen.

Diese fünf Merkmale sollen die minimalen Anforderungen eines Usability-Tests beschreiben. Jedoch ist es heute nicht mehr unbedingt notwendig, dass Teilnehmer während des Tests beobachtet werden, wie dies der im Kapitel 2.2.2 beschriebene unmoderierte Remote Test zeigt (Barnum, 2011: 13). Die vierte Anforderung könnte folgendermassen angepasst werden:

- 4b. Handlungen, Aussagen oder Meinungen der Teilnehmer werden festgehalten.

Dumas' und Redishs zweite Anforderung, tatsächliche Nutzer zu rekrutieren, ist umstritten. So stellt Krug (2006: 139 f.) fest, dass in den meisten Fällen irgendjemand an den Tests teilnehmen kann, falls er grundlegende Kenntnisse des Webs besitzt. Krug plädiert dafür, dass es sinnvoll aber nicht unbedingt nötig sei, tatsächliche Anwender zu rekrutieren. Eine Benutzerschnittstelle sollte nicht nur von der Zielgruppe genutzt werden können. Wenn etwas für Anfänger benutzerfreundlich sei, so Krug, sei es das auch für Experten.

Usability-Tests können auf unterschiedliche Art und Weise durchgeführt werden. Wenn Teilnehmer und Forscher sich am selben Ort, zum Beispiel in einem Usability-Labor oder in einem Büro befinden, ist von einem In-person Usability-Test die Rede. Nutzer-tests können aber auch über das Internet erfolgen. Dies sind so genannte Remote Usability-Tests. Diese zwei Methoden werden in den folgenden Kapiteln ausführlicher beschrieben.

2.2.1 In-person Usability-Test

Beim In-person Usability-Test befinden sich Teilnehmer und Moderator am selben Ort. Der Teilnehmer führt eine oder mehrere Aufgaben mit der zu testenden Benutzeroberfläche durch. Bei traditionellen Tests interveniert der Moderator nicht oder nur so wenig wie möglich und konzentriert sich auf das Beobachten. Er kann sich sogar in einem anderen Raum befinden und den Teilnehmer nur durch Kamera oder Einwegspiegel beobachten.

Immer öfter wird die Methode des lauten Denkens verwendet, bei welcher der Teilnehmer dazu angehalten wird, seine Gedanken laut zu formulieren. Dieses sogenannte Thinking Aloud soll erlauben, den Teilnehmer und sein Handeln besser zu verstehen. Da individuelles lautes Denken kein natürlicher Vorgang ist, ist es auch möglich, dass zwei Teilnehmer die Aufgaben gemeinsam lösen. Durch die Interaktion und die Gespräche soll eine unbefangene Atmosphäre geschaffen werden (Sarodnick, Brau, 2011: 170 f.).

Damit die Aufgaben später ausgewertet werden können, können Bildschirm, Körper, Ton und teilweise auch Augenbewegungen aufgezeichnet werden (Sarodnick, Brau, 2011: 168 f.). Der Vorteil dieser Methode ist, dass das Verhalten der Testpersonen sehr gut aufgezeichnet und bei Bedarf interveniert oder nachgefragt werden kann.

Der klassische Nutzertest wird in einem Labor durchgeführt, welches häufig aus einem Test- und eventuell einem Kontrollraum besteht. Die Testperson befindet sich im mit Mikrofon und Kameras ausgestatteten Testraum und wird vom Testleiter angeleitet und beobachtet. Teilweise sind noch Beobachter im Kontrollraum anwesend, welche den Test beobachten (Sarodnick, Brau, 2011: 168).

Häufig werden Usability-Tests im Labor nach strengen wissenschaftlichen Richtlinien geplant und durchgeführt. Daher gilt diese Methode als aufwendig und kostenintensiv (Schweibenz, Thissen, 2003: 130). Ausserdem wird kritisiert, dass durch die unnatürliche Umgebung und durch die Anwesenheit von Testleiter und Beobachter das Verhalten der Testpersonen beeinflusst werden kann.

Es gibt jedoch auch Labors, die mit wenig technischer Ausrüstung auskommen und dadurch preiswerter sind. Mobile Lösungen mit Laptop und Webkamera erlauben es heute, Usability-Tests an beliebigen Orten durchzuführen (Barnum, 2011: 37 f.). Dadurch wird ermöglicht, den In-person Usability-Test in der natürlichen Umgebung der Teilnehmer vorzunehmen. Damit der Aufwand eines In-person Tests reduziert wird, propagieren Nielsen⁹ und Krug (2006: 135 f.) die Durchführung von Discount Usability-Tests mit wenigen Teilnehmern und von kleinerem Umfang.

⁹ NIELSEN, Jakob. Discount usability: 20 years. In: *Nielsen Norman Group* [online]. 14.09.2009. <http://www.nngroup.com/articles/discount-usability-20-years/> (letzter Zugriff: 08.06.2013)

2.2.2 Remote Usability-Test

Der Remote Usability-Test ist eine Sonderform des In-person Usability-Tests und wurde vor ungefähr fünfzehn Jahren zum ersten Mal durchgeführt.¹⁰ Die Bezeichnung „remote“ (Englisch für „entfernt“) deutet bereits an, dass der Testleiter und die Probanden geografisch voneinander getrennt sind (Sarodnick, Brau, 2011: 179). Dabei wird der Test mittels Informations- und Kommunikationstechnologien wie dem Internet oder dem Telefon durchgeführt (Bolt, Tulathimutte, 2010: VI.). Es müssen sich somit weder Testleiter noch Studienteilnehmer anderswo hinbegeben und der Usability-Test kann in der gewohnten Umgebung stattfinden (Sarodnick, Brau, 2011: 179).

Es lassen sich zwei Formen unterscheiden: der moderierte und der unmoderierte Test. Beim moderierten Remote Usability-Test beobachtet ein Moderator den Teilnehmer, der sich an einem anderen Ort befindet, und kommuniziert mit diesem. Diese Methode wird auch synchroner Remote Usability-Test genannt, da Testleiter und Teilnehmer in Echtzeit kommunizieren. Da meistens wenige Teilnehmer rekrutiert werden, welche viele subjektive Rückmeldungen liefern, erfasst diese Methode vor allem qualitative Daten (Bolt, Tulathimutte, 2010: 22 f.). Bei einem moderierten Nutzertest wird der Bildschirm des Teilnehmers mit einer Software via Internet zum Moderator übertragen (Screen-Sharing) und teilweise auch aufgenommen. In seltenen Fällen wird der Teilnehmer zusätzlich mit einer Webcam aufgenommen. Teilnehmer und Moderator kommunizieren über eine Chatfunktion, über das Telefon oder mittels Internet-Telefonie, wobei der Dialog in manchen Fällen für die spätere Analyse gespeichert wird (Bolt, Tulathimutte, 2010: 32 ff.).

Beim unmoderierten Remote Usability-Test sind Testleiter und Teilnehmer nicht nur geografisch getrennt, sondern haben auch keinen direkten Kontakt. Teilweise werden auch die Begriffe asynchroner oder automatischer Remote Usability-Test verwendet. Zunächst wird der Test auf einer speziellen Software konzipiert. Diese Phase ist sehr wichtig, da während des Tests keine Interaktion mit dem Teilnehmer geschieht und nicht auf Unvorhergesehenes reagiert werden kann. Nach der Veröffentlichung des Tests können die Teilnehmer auf den Test zugreifen und selbstständig die Aufgaben durchführen. Während des Tests wird das Nutzerverhalten mit einer speziellen Software automatisch aufgezeichnet (Bolt, Tulathimutte, 2010: 22 f., 122 ff.).

¹⁰ ANDREASEN, Morten Sieker et al. What happend to Remote usability testing?: an emprical study of three methods. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 07)* [online]. New York: ACM, 2007. S. 1405-1414.
http://dl.acm.org/ft_gateway.cfm?id=1240838&ftid=416478&dwn=1&CFID=336910444&CFTOKEN=59384153 (letzter Zugriff: 08.06.2013)

Je nach verwendetem Programm werden folgende Daten erfasst (Albert, Tullis, Tedesco, 2010: 24 ff.):

- Erfolgsraten,
- Aufgabenzeiten,
- Klickstream-Daten,¹¹
- Mausklicks auf einer Seite (Heatmaps),
- Anzahl besuchter Seiten,
- oder andere.

Einige Tools zeichnen zusätzlich Bildschirm oder Stimme auf, diese bilden jedoch die Ausnahme.¹² Üblicherweise werden solche Tests mit fünfzig oder mehr Teilnehmern durchgeführt und liefern daher eher quantitative Daten, die statistisch ausgewertet werden können. Da es zu keiner direkten Interaktion zwischen Moderator und Teilnehmer kommt, ist schwierig herauszufinden, was die Beweggründe der Benutzer sind und weshalb sie etwas tun. Deshalb werden häufig im Anschluss an die Aufgaben Fragen oder Kommentarmöglichkeiten eingefügt, die es erlauben, zusätzliche Informationen und subjektive Rückmeldungen zu erhalten (Bolt, Tulathimutte, 2010: 22 f., 122 ff.).

2.2.3 Discount Usability-Test vs. unmoderierter Remote Usability-Test

Verschiedene Studien beschäftigen sich mit dem Unterschied zwischen In-person und Remote Usability-Tests (Andreasen et al. 2007, Bruun 2009, Kelly, Gyllstrom, 2011, Meier, Wolff, 2012, Selvaraj, 2004, Tullis et al. 2002, Waterson, Landay, 2002, West, Lehmann, 2006). Keine dieser Arbeiten vergleicht aber die zwei in dieser Arbeit durchgeführten Evaluationsmethoden *Discount Usability-Test* und *unmoderierter Remote Usability-Test*. Die Resultate dieser Studien werden deshalb nicht näher erläutert. In diesem Kapitel wird vor allem auf die allgemeinen Unterschiede zwischen einem Remote und einem In-person Test eingegangen und diese werden auf den Discount und den unmoderierten Remote Tests angepasst.

Ein Remote Usability-Test bietet gegenüber der traditionellen In-person Forschung verschiedene Vorteile. Der Studienteilnehmer kann den Test eigenständig und zu einer beliebigen Zeit durchführen. Er befindet sich dabei in seiner natürlichen Umgebung, wodurch die Handlungen weniger beeinflusst werden sollen (Sarodnick, Brau, 2011: 180 f.). Jedoch kann ein Discount Usability-Test ebenfalls an einem beliebigen Ort

¹¹ Klickstream-Daten zeigen auf, wie Benutzer durch die Webseite navigiert haben.

¹² ETHNIO. Remote usability and UX research tools. In: *Remote usability testing* [online]. 2013. <http://remoteresea.ch/tools> (letzter Zugriff: 17.06.2013)

stattfinden, da kein ausgestattetes Usability-Labor zur Durchführung benötigt wird. Die Beeinflussung der Teilnehmer durch die Umgebung kann daher verkleinert werden.

Bei einem Remote Test ist es einfacher, Teilnehmer aus unterschiedlichen Regionen zu rekrutieren, die bei einem In-person Test schwer erreichbar sind. Der automatisierte Ablauf erlaubt es, eine grosse Stichprobe zu erheben und somit eine repräsentativere Umfrage zu ermöglichen. Ein unmoderierter Usability-Test eignet sich daher besonders für quantitative Fragestellungen (Sarodnick, Brau, 2011: 180 f.). Für einen Discount Nutzertests werden durchschnittlich fünf Personen getestet, weshalb die Aussagekraft der erhobenen statistischen Daten in Frage gestellt werden kann. Auch wenn Remote Nutzertests gut sind, um geografisch schwierig erreichbare Nutzergruppen zu rekrutieren, muss bedacht werden, dass nicht alle Menschen über eine zuverlässige Internetverbindung verfügen (Albert, Tullis, Tedesco, 2010: 10). Jedoch ist dieser Nachteil für Remote Tests von Internetseiten zu vernachlässigen, da diese Personen nicht der Nutzergruppe entsprechen.

Beim Discount Test können die Teilnehmer beobachtet werden und bei Unklarheiten kann nachgefragt werden. Dadurch werden viele qualitative Daten gesammelt. Direktes Beobachten ist mit einem unmoderierten Remote Usability-Test nicht möglich, ausser es werden zusätzliche Video- oder Tonaufnahmen gemacht.¹³ Üblicherweise kann aber die Mimik und Gestik des Benutzers nicht beobachtet werden. Durch das Einfügen von Kommentarmöglichkeiten oder offenen Fragen ist es möglich, diesen Nachteil zu verkleinern, da so zusätzliche Informationen über den Kontext erfasst werden können. Solche überlegten Kommentare haben jedoch nicht den gleichen Wert wie spontane Gedankenäusserungen und es ist nicht möglich, nachzufragen und daher schwierig herauszufinden, aus welchen Beweggründen ein Teilnehmer agiert (Bolt, Tulathimutte, 2010: 123 f.).

Ein weiterer Nachteil des Remote Tests ist seine Unflexibilität. Es können nicht beliebig komplexe Aufgaben konzipiert werden, da die Umsetzung je nach verwendetem Tool nicht möglich ist. Bei zu schwierigen Aufgaben würden auch mehr Teilnehmer den Test einfach abbrechen (Barnum, 2011: 44 f.). Die Aufgaben müssen sehr gut durchdacht sein, da nicht auf unerwartete Ereignisse reagiert werden kann (Sarodnick, Brau, 2011: 180). Damit die Abbruchrate niedrig bleibt, ist auch der Umfang einer Online Studie klein zu halten, denn bereits nach 20 bis 30 Minuten sinkt die Aufmerksamkeit der Teilnehmer (Albert, Tullis, Tedesco, 2010: 10).

¹³ SAURO, Jeff. 10 things to know about unmoderated usability testing. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 13.11.2012.
<http://www.measuringusability.com/blog/unmoderated-things.php> (letzter Zugriff: 08.06.2013)

Ein wichtiger Aspekt für die Wahl einer bestimmten Evaluationsmethode sind häufig die Kosten. Bei einem Discount Usability-Test wird versucht, diese so tief wie möglich zu halten. Daher werden einfache und kostengünstige Mittel gewählt. Bei Remote Tests wird häufig argumentiert, dass die Ausgaben durch wegfallende Transport oder Laborkosten minimiert werden können.¹⁴ Bolt und Tulathimutte (2010: 239) und auch Barnum (2011: 45) widersprechen jedoch dieser Annahme. Den Grund dafür sehen sie in der aufwendigen Analyse der Daten, welche einen grossen Teil der Kosten ausmacht.

Bolt und Tulathimutte (2010: 6 f.) weisen noch darauf hin, dass Remote Nutzertests nicht dieselbe Sicherheit bieten können wie In-person Tests. Falls das Produkt sensible Daten beinhaltet, die nicht veröffentlicht werden sollen oder falls es sich um ein geheimes Projekt handelt, ist es sicherer, einen In-person Test durchzuführen. Zudem kann es bei einem Remote Test zu Schwierigkeiten mit der Technik kommen und teilweise wird auch die Installation von Software auf dem Computer der Teilnehmer benötigt.

¹⁴ GOUGH, Dabney, PHILLIPS, Holly. Remote online usability testing: why, how, and when to use it. In: *Boxes and arrows* [online]. 09.06.2003.
http://www.bboxesandarrows.com/view/Remote_online_usability_testing_why_how_and_when_to_use_it
(letzter Zugriff: 05.07.2013)

3 Forschungsfragen

Um die zwei Evaluationsmethoden zu vergleichen, wurden drei Hypothesen über die Unterschiede eines unmoderierten Remote und eines Discount Usability-Tests verfasst. Die Aufgaben der Tests wurden so gestaltet, dass sie möglichst viele Daten über eventuelle Differenzen liefern. Durch die gewonnenen Daten und durch den Vergleich der beiden Evaluationsmethoden sollen die Hypothesen falsifiziert oder gestützt werden.

3.1 Subjektive Einstellung der Teilnehmer

Einer der zu vergleichenden Punkte war die subjektive Einstellung der Teilnehmer während eines Tests. Subjektive Einstellungen werden bei Usability-Tests oft durch Fragen erhoben. Jedoch bemerkt Nielsen¹⁵, dass Nutzer oft nicht ehrlich sind, sondern dasjenige antworten, was sie denken, dass es gehört werden will oder was sozial akzeptabel ist. Bezieht sich diese Aussage nur auf In-person Tests oder agieren Nutzer bei einem Remote Test genauso? Es ist auch möglich, dass die Testmethode die Teilnehmer in ihren Antworten beeinflusst. Beim Remote Test befindet sich der Teilnehmer in seiner gewohnten Umgebung. Er kann anonym arbeiten und wird nicht von jemandem beobachtet. Beim In-person Test hingegen kommt es zu einer Interaktion zwischen Moderator und Teilnehmer und der Teilnehmer befindet sich an einem unbekannten Ort in einer ungewohnten Situation. Folgende Hypothese wurde zu dieser Fragestellung verfasst:

Hypothese 1: *Die subjektive Einstellung der Testpersonen während eines unmoderierten Remote Usability-Tests ist nicht identisch mit der Einstellung während eines In-person Usability-Tests, da die Umgebungen die Testpersonen unterschiedlich beeinflussen.*

Diese Hypothese sollte durch verschiedene Indikatoren untersucht werden, welche unter dem Kapitel 6.1 näher erläutert werden.

- Single Ease Question (SEQ) am Ende jeder Aufgabe
- System Usability Scale (SUS) gegen Ende des Tests
- Zusätzliche Fragen zum Testverlauf am Ende des Tests

Die SEQ soll es erlauben, zu vergleichen, wie schwierig die Aufgaben je nach Testmethode wahrgenommen werden. Durch die standardisierte SUS wird gemessen, wie Benutzer die allgemeine Benutzerfreundlichkeit eines Produkts wahrnehmen. Sechs zusätzliche Fragen am Ende des Tests sollen evaluieren, wie sich die Probanden während des Tests gefühlt haben und wie sie zu der jeweiligen Methode stehen.

¹⁵ NIELSEN, Jakob. First rule of usability?: Don't listen to users. In: *Nielsen Norman Group* [online]. 05.08.2001. <http://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/> (Zuletzt besucht: 20.06.2013)

3.2 Verhalten der Teilnehmer

Eine weitere Forschungsfrage beschäftigt sich mit den Auswirkungen eines Usability-Tests auf das Verhalten. Albert, Tullis und Tedesco (2010: 8) verweisen darauf, dass Teilnehmer einer Online Studie tiefere Erfolgsraten oder längere Aufgabenzeiten haben können, da sie teilweise über kleinere Bildschirme verfügen, schlechtere Auflösungen haben oder unterschiedliche Browser nutzen. Neben technischen Aspekten können Aufgabenzeiten auch durch die Ablenkungen aus dem Alltag verlängert werden. Im Gegensatz dazu konzentrieren sich Teilnehmer während eines In-person Tests ausschliesslich auf die Aufgaben (Albert, Tullis, Tedesco, 2010: 8). Denkbar wäre aber auch, dass sich die Teilnehmer eines In-person Tests mehr Zeit lassen und genauer suchen, da sie beobachtet werden und die Aufgaben gut und richtig lösen wollen. Mangel an Konzentration und Geduld könnten während eines Remote Tests zu früherem Beenden einer Aufgabe führen, was mitunter auch die Erfolgsrate und die Anzahl besuchter Seiten beeinflussen würde. Die Hypothese 2 soll dabei helfen, diese Vermutungen zu untersuchen:

Hypothese 2: *Das Verhalten einer Testperson während eines unmoderierten Remote Usability-Tests ist anders und realistischer als während eines klassischen Nutzertests.*

Folgende drei Indikatoren dienen der Überprüfung der Hypothesen:

- Zeit pro Aufgabe
- Besuchte Seiten pro Aufgabe
- Erfolgsrate

Loop11 generiert automatisch die Zeiten und die Anzahl besuchter Seiten pro Aufgabe und Teilnehmer. Für den In-person Test wurden diese manuell durch das Visionieren der Videoaufnahmen ermittelt. Über Erfolg oder Misserfolg wurde anhand der Richtigkeit der Multiple-Choice-Fragen am Ende jeder Aufgabe entschieden, wobei es für Aufgaben 1 und 6 keine richtige oder falsche Antwort gab, da diese Fragen eher qualitative Daten lieferten (siehe Anhang C).

3.3 Effektivität der Methoden

Effektivität hilft zu bestimmen, wie gut ein festgelegtes Ziel erreicht wurde. Dieses Mass ist unabhängig vom Aufwand, der betrieben wird, um das Ziel zu erreichen. Das primäre Ziel beim Durchführen eines Usability-Tests ist es, Probleme zu identifizieren, um Verbesserungen für diese vorzuschlagen (Sarodnick, Brau, 2011: 163). Damit eine Usability-Test-Methode als effektiv gilt, muss sie daher möglichst gut Usability-Probleme identifizieren können. Hierbei ist es nicht nur wichtig, dass viele Probleme identifiziert werden, sondern auch, wie gross die Wirkung der jeweiligen Probleme ist und inwiefern sie die Benutzer in ihrem Handeln beeinflussen (Sarodnick, Brau, 2011: 198). Als Untersuchungsgrundlage diene folgende Hypothese:

Hypothese 3: *Ein unmoderierter Remote Usability-Test ist genauso effektiv wie ein In-person Discount Usability-Test.*

Um zu überprüfen, wie effektiv ein asynchroner Remote Usability-Test im Vergleich zu einem Discount Usability-Test ist, werden zwei Indikatoren verwendet:

- Anzahl identifizierter Usability-Probleme
- Schweregrad der identifizierten Usability-Probleme

Die Anzahl der Nutzungsprobleme lässt sich relativ einfach feststellen. Schwieriger wird es hingegen bei der Zuordnung eines Schweregrads. Nielsen und Loranger (2006: 125) beschreiben drei Faktoren, welche die Schwere eines Problems beeinflussen:

- die Häufigkeit (Bei wie vielen Nutzern tritt das Problem auf?)
- die Auswirkung (Wie sehr stört das Problem den Nutzer; kann er das Problem einfach lösen oder nicht?)
- die Beständigkeit (Können Nutzer lernen, ein Problem zu überwinden oder tritt es immer wieder auf?)

Es gibt unterschiedliche Skalen mit mehr oder weniger Stufen, die Usability-Probleme gewichten. In Anlehnung an Barnum (2011: 264) wird eine sehr einfache Skala mit drei unterschiedlichen Schweregraden genutzt: schwerwiegend, wesentlich oder marginal. Diese wird im Kapitel 5 näher beschrieben.

4 Durchführung der Nutzertests

Der In-person Usability-Test und der unmoderierte Remote Usability-Test wurden verglichen, indem dasselbe Testszenario mit beiden Methoden durchgeführt wurde. Bis auf einige wenige Ausnahmen, die in den folgenden Kapiteln beschrieben werden, wurden die Testabläufe so identisch wie möglich konzipiert. Als Testobjekt wurde die fertige Webseite Swissbib gewählt (siehe Kapitel 1.2). Es handelte sich somit um einen summativen Nutzertest. Es wurde ein Test von kleinem Umfang konzipiert. Dies diente dazu, die Abbruchrate beim Remote Nutzertest niedrig zu halten und entspricht zudem dem Konzept eines Discount Nutzertests (siehe Kapitel 2.1.2). Die Unterlagen, welche für den konkreten Testablauf benutzt wurden, befinden sich in den Anhängen A bis F.

4.1 Konzeption des Tests

4.1.1 Pretest-Fragebogen

Der Pretest-Fragebogen wird vor den eigentlichen Aufgaben von den Teilnehmern ausgefüllt und liefert zusätzliche relevante Informationen über jeden Teilnehmer (Barnum, 2011: 173). Um das Auswerten zu erleichtern, wurden nur geschlossene Fragen mit Ankreuzmöglichkeiten gestellt. Für statistische Zwecke wurde zuerst nach Alter und Geschlecht gefragt. Auf weitere demografische Fragen wurde verzichtet, um zu persönliche oder zu heikle Fragen zu verhindern und die Teilnehmer dadurch nicht zu verunsichern. Aus diesem Grund wurde auch nicht nach dem genauen Alter, sondern nach der Altersgruppe gefragt (Albert, Tullis, Tedesco, 2010: 87 f.). Ausserdem wurde nach der Erfahrung mit dem Internet, Bibliothekskatalogen und Swissbib gefragt, da diese Faktoren die Geschwindigkeit und den Erfolg beim Lösen der Aufgaben beeinflussen können (Albert, Tullis, Tedesco, 2010: 59 f.).

4.1.2 Aufgaben und Szenarien

Das primäre Ziel der Nutzertests war es, zwei unterschiedliche Evaluationsmethoden zu vergleichen. Daher wurden einfache und typische Aufgaben gewählt, die erlauben, mögliche Unterschiede im Verhalten und in den Einstellungen während den jeweiligen Tests zu identifizieren. Die Aufgaben wurden somit nicht aus potentiellen Usability-Problemen, aus eigenen Bedenken oder Erfahrungen oder aus anderen Kriterien abgeleitet, was ebenfalls häufige Verfahren sind (Schweibenz, Thissen, 2003: 140). Insgesamt wurden sechs Aufgaben gewählt, welche nach Schwierigkeitsgrad geordnet wurden, beginnend mit der einfachsten (siehe Anhang C).

Zuerst wurde der Teilnehmer dazu aufgefordert, ein beliebiges Buch zu suchen, um sich mit der Webseite vertraut zu machen und um erste Eindrücke festzuhalten. Auf diesen Einstieg folgten vier Aufgaben, bei welchen die Benutzer eine spezifische Information auf Swissbib suchen mussten. Diese Fragen hatten eine eindeutige Antwort, welche durch eine Multiple-Choice-Frage geprüft wurde (siehe Kapitel 4.1.3). Die sechste und letzte Aufgabe bestand darin, sich auf Swissbib anzumelden und herauszufinden, welche Vorteile ein Benutzerkonto bietet. Da viele unterschiedliche Funktionen entdeckt werden konnten, gab es keine eindeutig richtige oder falsche Antwort. Das Ziel dieser Aufgabe war eher, zu unterscheiden, ob In-person oder Remote Teilnehmer mehr Zeit investierten, um Informationen zu finden.

Damit die Aufgaben für die Teilnehmer so realitätsnah wie möglich erschienen, wurden Szenarien formuliert. Diese stellten eine mögliche Nutzersituation dar, in welche sich der Teilnehmer hineinversetzen und aus welcher heraus er dann mit der Benutzeroberfläche interagieren sollte. Die Szenarien wurden gemäss den Empfehlungen von Dumas und Redish (1999: 173 ff.) kurz und einfach gehalten und in der Sprache der Benutzer formuliert. Um zu gewährleisten, dass die Aufgaben nicht mehrdeutig interpretiert werden konnten, wurde der Test vorgängig von mehreren Personen erprobt.

4.1.3 Folgefragen zu den Aufgaben

Üblicherweise werden in einem traditionellen Usability-Test keine Fragen nach jeder Aufgabe gestellt, da eine möglichst natürliche Nutzungssituation simuliert wird (Dumas, Redish, 1999: 210). Jedoch ist dies bei Remote Studien nützlich, da nur so zusätzliche Informationen über den Kontext gewonnen werden können. Auf offene Fragen wurde bis auf zwei Ausnahmen (Aufgabe 1 und 6) verzichtet, um die Auswertung der Daten zu vereinfachen (Albert, Tullis, Tedesco, 2010: 59 f.).

Im Anschluss an die Aufgaben 2 bis 5 wurde mit einer Multiple-Choice-Frage nach der genauen Antwort auf die Aufgabe gefragt. Dieses untypische Verfahren war notwendig, da die von Swissbib generierten Links zu lange waren, um als Erfolgslink definiert zu werden. Durch die Multiple-Choice-Frage sollte eindeutig festgestellt werden können, ob die Teilnehmer die Aufgabe lösen konnten oder nicht. Anschliessend wurde die Single Ease Question verwendet, um die Teilnehmer nach jeder Aufgabe zu fragen, wie einfach respektive schwierig die Aufgabe für sie war. Die Skala reicht von 1 für sehr schwierig bis 7 für sehr einfach (siehe Abb. 1).¹⁶

¹⁶ SAURO, Jeff. 10 things to know about the single ease question (Seq) In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 30.10.2012.
<http://www.measuringusability.com/blog/seq10.php> (letzter Zugriff: 20.06.2013)

Abbildung 1: Single Ease Question

Wie schwierig oder einfach fanden Sie diese Aufgabe?								
	1	2	3	4	5	6	7	
Sehr schwierig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sehr einfach

Diese Frage ist kurz und einfach auszuwerten und kann sowohl für Remote als auch In-person Studien verwendet werden. Die Verwendung dieser Skala zeigt wie Teilnehmer eine Aufgabe einschätzen und wie zufrieden sie damit waren.¹⁷ So können durch den Vergleich der Resultate eventuelle Unterschiede in den Einstellungen der Teilnehmer während beider Tests identifiziert werden (vgl. Hypothese 1, Kapitel 3.1).

4.1.4 Posttest-Fragebogen

Nach dem Lösen aller Aufgaben füllte jeder Teilnehmer zwei Posttest-Fragebögen aus. Verwendet wurde die System Usability Scale (SUS), die auf Deutsch übersetzt wurde. Dieser standardisierte Fragebogen enthält zehn Fragen mit einer Fünf-Punkte Skala (siehe Abb. 2).

Abbildung 2: Beispielfrage der System Usability Scale

3. Ich fand, Swissbib war einfach zu benutzen.						
	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

Die System Usability Scale wurde wegen ihrer Verbreitung, ihrer Kürze, ihrer Einfachheit und ihrer freien Verfügbarkeit gewählt. Die SUS soll zudem auch bei einer kleinen Anzahl Teilnehmern verlässliche Resultate liefern, was im Hinblick auf die fünf Teilnehmer beim In-person Usability-Test ein wichtiger Faktor war (Albert, Tullis, Tedesco, 2010: 85 f.).

Neben der SUS wurde noch ein weiterer Fragebogen mit sechs zusätzlichen Fragen erstellt, der von einer ähnlichen Studie inspiriert wurde (Selvaraj, 2004: 61). Diese Fragen bezogen sich auf die Testmethode und sollten evaluieren, wie sich die Teilnehmer gefühlt haben und wie sie zu dem Test stehen (vgl. Hypothese 1, Kapitel 3.1).

¹⁷ SAURO, Jeff. If you could only ask one question, use this one. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 30.10.2012. <http://www.measuringusability.com/blog/single-question.php> (letzter Zugriff: 02.03.2010)

4.2 Umsetzung des Tests – In-person

4.2.1 Datenerhebung

Für die Umsetzung des In-person Tests wurde die Software Morae¹⁸ von TechSmith verwendet, welche Daten für einen Usability-Test erhebt und auswertet. Mit der Software wurden Mausbewegungen, Mausklicks und Ton aufgezeichnet. Alle Teilnehmer gaben ihr Einverständnis, zusätzlich mit einer Webkamera gefilmt zu werden. Zudem wurde die Thinking Aloud Methode verwendet (siehe Kapitel 2.2.1). Mit Fragen wie:

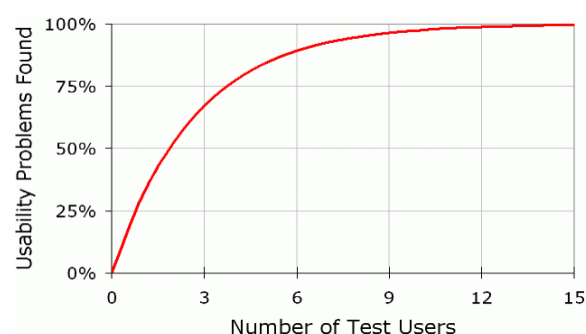
- Wonach suchen Sie jetzt?
- Was denken Sie, hat das zu bedeuten?
- Was haben Sie erwartet, würde passieren?

wurden die Teilnehmer daran erinnert, laut zu denken (Sarodnick, Brau, 2011: 171, Barnum, 2011: 210 ff.).

4.2.2 Vorbereitungen

Es ist umstritten, wie viele Teilnehmer für einen In-person Usability-Test rekrutiert werden sollten.¹⁹ Gemäss den Empfehlungen von Nielsen²⁰, dem Erfinder des Discount Usability-Tests, ist es besser häufig aber dafür mit wenigen Teilnehmern zu testen. Beim In-person Test wurden fünf Teilnehmer rekrutiert, da diese ungefähr 80% der Usability-Probleme identifizieren sollen (Abb. 3). Es wurden fünf Personen aus dem persönlichen Umfeld rekrutiert.

Abbildung 3: Identifizierte Usability-Probleme nach Anzahl Teilnehmer



(Quelle: www.nngroup.com)

¹⁸ TECHSMITH. Morae. In: *TechSmith: wir machen Snagit, Camtasia, Jing und andere Software* [online]. 2013. <http://www.techsmith.de/morae.html> (letzter Zugriff: 08.06.2013)

¹⁹ BARNUM, Carol et al. The „magic number 5“: is it enough for web testing? In: *Extended abstracts on human factors in computing systems (CHI EA '03)* [online]. New York: ACM, 2003. S. 698-699 http://dl.acm.org/ft_gateway.cfm?id=765936&ftid=152053&dwn=1&CFID=336910444&CFTOKEN=59384153 (letzter Zugriff: 08.06.2013)

²⁰ NIELSEN, Jakob. Why you only need to test with 5 users. In: *Nielsen Norman Group* [online]. 19.03.2000. <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/> (letzter Zugriff: 08.06.2013)

Bei einem Discount Usability-Test ist es weniger wichtig mit Personen zu testen, die auch der tatsächlichen Zielgruppe entsprechen (Krug, 2006: 140). Daher wurde zwar versucht mit Nutzern wie Studenten oder Bibliothekaren zu testen, jedoch entsprachen die Teilnehmer 2 und 5 nicht der primären Nutzergruppe von Swissbib, die eher aus Studenten, Hochschulangehörigen oder Informationsspezialisten besteht (siehe Tabelle 2). Es wurde versucht, unterschiedliche Altersgruppen, Erfahrungen und Geschlechter zu berücksichtigen.

Tabelle 2: Teilnehmer des In-person Tests

	<i>Geschlecht</i>	<i>Alter</i>	<i>Beruf</i>	<i>Erfahrung Internet</i>	<i>Erfahrung Bibliotheks-kataloge</i>	<i>Erfahrung Swissbib</i>
1	weiblich	21-30	Studentin (Heilpädagogik)	4	3	1
2	männlich	21-30	Architekt	4	2	1
3	männlich	31-40	Bibliothekar	4	5	4
4	weiblich	21-30	Studentin (Germanistik)	4	4	3
5	weiblich	51-60	Sozialpädagogin	2	1	1

Nach der Rekrutierungsphase wurden gemäss den Empfehlungen von Barnum (2011: 163 ff.) unterschiedliche Dokumente erstellt. Für jeden Teilnehmer wurde ein Skript erstellt, welches eine Einverständniserklärung (Anhang E), die Aufgaben (Anhang C) und Fragebögen (Anhang B und D) enthielt. Die Einverständniserklärung informierte die Teilnehmer über den Zweck der Studie und die Verwendung der erhobenen Daten. Die Teilnehmer wurden auf ihr Recht hingewiesen, einzelne Aufgaben oder den gesamten Test jederzeit abbrechen zu können und sie hatten die Möglichkeit, auf die Aufnahme mit einer Kamera zu verzichten. Zusätzlich wurde ein Skript für den Testleiter erstellt, das ihn während des Tests leiten sollte (siehe Anhang A). Es enthielt Listen mit zu erledigenden Aufgaben und einen Text mit den Informationen, die jedem Teilnehmer mitgeteilt werden sollten. Nachdem der Test konzipiert war, wurde er mit zwei Experten für Usability getestet und gemäss ihren Rückmeldungen angepasst.

4.2.3 Testablauf

Die Software Morae war auf einem Laptop installiert. Dies erlaubte den mobilen Einsatz, wie es für Discount Tests üblich ist (Krug, 2006: 137). Alle fünf Tests wurden somit nicht in einem Usability-Labor, sondern in einem Büro in Leuk-Stadt (Kanton Wallis) durchgeführt.

Jeder Teilnehmer wurde gemäss dem Skript für den Testleiter begrüsst und instruiert. Nach dem Ausfüllen der Einverständniserklärung wurden ihm die Aufgaben und Fragebögen ausgehändigt. Der Teilnehmer bearbeitete die Aufgaben mit dem zur Verfügung gestellten Laptop und wurde dazu angehalten, laut zu denken. Sobald der Test beendet war, wurde dem Teilnehmer gedankt und er wurde mit einem Geschenkgutschein verabschiedet.

4.3 Umsetzung des Tests – Remote

4.3.1 Datenerhebung

Es existieren viele Tools für die Konzeption eines unmoderierten Usability-Tests. Eine umfangreiche Liste findet sich unter <http://Remotereseach.ch/tools/>. Für die vorliegende Studie wurde der Online Dienst Loop11²¹ gewählt. Loop11 wurde genutzt, da es eines der einzigen Tools ist, welches erlaubt, Tests ohne Installation einer Software und bis auf wenige Ausnahmen auch ohne Veränderungen im Quellcode der Webseite durchzuführen. Ausserdem bietet Loop11 eine kostenlose Version für Studenten im Bereich Usability. Mit diesem Tool können Aufgaben und Fragebögen zu jeder beliebigen Webseite erstellt werden, welche jeder Teilnehmer autonom durchläuft. Verschiedene Funktionen wie Berechnung der Erfolgsraten und Zeit pro Aufgabe oder die Analyse der Navigationspfade werten das Verhalten der Teilnehmer automatisch aus. Die Resultate können auch in verschiedenen Formaten exportiert werden.²²

4.3.2 Vorbereitungen

Die Aufgaben und Fragebögen, die im Anhang B bis D zu finden sind, wurden auf Loop11 umgesetzt. Zusätzlich war es nötig, einen Begrüssungs- und einen Verabschiedungstext zu verfassen. Beispiele für die Umsetzung des Nutzertests auf Loop11 befinden sich im Anhang G. Vor der Lancierung des Tests wurde dieser mit sieben Studenten der Informationswissenschaften und zwei Usability-Experten getestet.

²¹ LOOP11. *Loop11: Remote and online usability testing tool* [online]. 2013. <http://www.loop11.com/> (letzter Zugriff: 08.06.2013)

²² LOOP11. Features. In: *Loop11: Remote and online usability testing tool* [online]. 2013. <http://www.loop11.com/features/> (letzter Zugriff: 17.06.2013)

Nach dem Testlauf wurden auf Grund unterschiedlicher Rückmeldungen noch letzte Anpassungen vorgenommen.

Um Teilnehmer zu rekrutieren, wurde ein Link zum Test generiert. Dieser wurde in einer Mail an die Abonnenten von Swiss-Lib²³ verschickt (siehe Anhang F). In der E-Mail wurde das Ziel der Studie beschrieben und kurz erklärt, was Swissbib ist. Um möglichst viele zum Mitmachen anzuregen, wurde darauf hingewiesen, dass der Test zu jeder Zeit und von überall durchgeführt werden könne, dass er bloss 15 bis 20 Minuten dauern würde und dass jeder Teilnehmer ein kleines Dankeschön erhalten würde.

Da Swiss-Lib eine Schweizer Diskussionsliste für Spezialisten im Bereich Information und Dokumentation wie Bibliothekare, Archivare oder Dokumentalisten ist, konnte durch diese Rekrutierungsmethode eine spezifische Nutzergruppe von Swissbib angesprochen werden. Ein Grossteil der Teilnehmer war weiblich (Tabelle 4) und die Verteilung der Altersgruppen war sehr ausgeglichen (Tabelle 3). Einzig Personen unter 20 oder über 61 Jahren wurden nicht erreicht. Dies liegt wohl daran, dass wenige Leute dieser Altersgruppen Swiss-Lib abonniert haben. Swiss-Lib ist vor allem für Personen interessant, die sich bereits oder noch im Berufsleben befinden. Ein weiterer Grund für die Unterrepräsentation der über 60-Jährigen ist die weniger starke Internetnutzung dieser Altersgruppe.²⁴

Tabelle 3: Alter der Teilnehmer des Remote Tests

Alter	Verteilung
bis 20	3%
21 bis 30	27%
31 bis 40	23%
41 bis 50	20%
51 bis 60	22%
+61	5%

Tabelle 4: Geschlecht der Teilnehmer des Remote Tests

Geschlecht	Verteilung
Weiblich	70%
Männlich	30%

²³ SWISS-LIB. *Swiss-Lib: Schweizerische Diskussionsliste der SpezialistInnen für Information und Dokumentation* [online]. <http://lists.switch.ch/mailman/listinfo/swiss-lib> (letzter Zugriff: 17.06.2013)

²⁴ BUNDESAMT FÜR STATISTIK. Internetnutzung. In: *Statistik Schweiz* [online]. 2013. http://www.bfs.admin.ch/bfs/portal/de/index/themen/16/04/key/approche_globale.indicator.30106.301.html?open=302,5#5 (letzter Zugriff: 20.06.2013)

Die Tabelle 5 zeigt die Verteilung der Erfahrungen der 101 Teilnehmer. Entsprechend der Rekrutierungsmethode waren fast alle Teilnehmer erfahren, sehr erfahren oder gar Experten im Umgang mit Bibliothekskatalogen.

Tabelle 5: Erfahrungen der Teilnehmer des Remote Tests

	<i>Gar nicht erfahren</i>	<i>Wenig erfahren</i>	<i>Erfahren</i>	<i>Sehr erfahren</i>	<i>Experte</i>
<i>Erfahrung Internet</i>	0%	0%	24%	56%	20%
<i>Erfahrung Bibliotheks- kataloge</i>	0%	1%	23%	49%	27%
<i>Erfahrung Swissbib</i>	10%	19%	41%	28%	2%

4.3.3 Testablauf

Jedem Teilnehmer, der den via Swiss-Lib verschickten Link aufrief, wurde als erstes ein Begrüssungstext angezeigt. In diesem wurde erneut darauf hingewiesen, wie die Daten verwendet würden und dass alle Teilnehmer ein kleines Dankeschön erhielten. Um den Ablauf des Tests zu erklären, wurde die Standard-Erklärung von Loop11 verwendet (siehe Anhang G). Anschliessend durchlief jeder Teilnehmer die definierten Aufgaben und Fragen. Am Ende des Tests wurde ihnen gedankt und sie wurden aufgefordert, eine E-Mail an den Testleiter zu schicken, um ihr Dankeschön zu erhalten. So konnte gewährleistet werden, dass die Teilnehmer frei entscheiden konnten, ob sie ihre E-Mail-Adresse preisgeben wollten oder nicht.

Die Limite für die Teilnehmerzahl wurde bei hundert festgelegt. Bereits nach einer Woche hatten 101 Personen teilgenommen und der Test wurde daraufhin geschlossen.

5 Usability-Probleme Swissbib

Die Aussage eines Teilnehmers des Remote Tests sagt viel über die Benutzerfreundlichkeit von Swissbib aus: „Swissbib ist ein gutes, aber weiterhin verbesserungswürdiges Produkt.“ Generell wurde die Usability im durchgeführten Test von den Teilnehmern sehr positiv eingeschätzt. Ein durchschnittlicher Wert von fast 80 von möglichen 100 Punkten bei der System Usability Scale ist sehr hoch. Dieser sehr hohe Wert darf aber nicht falsch interpretiert werden. Die Aufgaben des Tests waren sehr einfach gestaltet und haben nur die elementarsten Funktionen von Swissbib getestet. Dies relativiert diesen hohen Wert ein wenig. Trotz der guten Einschätzung der Benutzer haben die beiden Tests unterschiedliche Schwachstellen aufgezeigt, die noch verbessert werden können.

Bevor in den nächsten Kapiteln ausführlich auf die spezifischen Probleme eingegangen wird, sollen noch kurz einige generelle Feststellungen zu der Benutzeroberfläche von Swissbib dargelegt werden. Allgemein nahmen die Benutzer vor allem Texte als klein und schlecht lesbar wahr. Ein Teilnehmer des Online Tests kommentierte Swissbib mit folgender Aussage: „Etwas kleine Schrift, für Leute in meinem Alter (>60) mit Lesebrille ist sie an der Grenze.“ Ein anderer Nutzer meinte: „Alles ist eher klein und sticht nicht heraus.“ Die seitliche Navigation bei den Untermenüs wird ebenfalls kritisiert und ein Teilnehmer meinte: „Ich finde, der linke Teil ist sehr schlecht zu sehen.“ Dieses Problem wird auch ersichtlich bei der Benutzung der Fassettennavigation (siehe Kapitel 5.4, Problem Nr. 9). Schliesslich ist es ein grundsätzliches Problem für Benutzer, dass Swissbib nur als Suchoberfläche dient und für Bestellungen auf einen lokalen Katalog zurückgegriffen werden muss. Dieses Problem wird verdeutlicht durch die schlechte Erfolgsrate bei der Aufgabe 4 (siehe Kapitel 6.2.3), bei der Benutzer die Verfügbarkeit eines Dokuments überprüfen sollten. Ein Teilnehmer beschreibt dieses Problem sehr treffend:

„Das Problem bei Swissbib ist wahrscheinlich weniger Swissbib selber als die Tatsache, dass man in verschiedenen Verbänden sucht und sobald man mehr über ein bestimmtes Dokument wissen will, auf die lokalen Oberflächen mit ihren unterschiedlichen Erscheinungsbildern weitergeleitet wird, mehr noch, wenn man sich ein Dokument auch beschaffen möchte, dann fangen die wirklichen Probleme an.“

Swissbib plant für das Jahr 2013, dass die Bestellung von Dokumenten direkt auf dem Metakatalog möglich wird.²⁵

²⁵ SWISSBIB. Über swissbib. In: *Swissbib* [online].
<http://www.swissbib.ch/TouchPoint/aboutswissbib.do;jsessionid=CED706B2067574E94AC6ED1D974C688C.worker4?methodToCall=showAboutSwissbib&submenue=nav1&lang=de> (letzter Zugriff: 05.04.2013)

Damit die Probleme nach ihrem Schweregrad geordnet werden konnten, wurde eine Ampel mit drei Stufen erstellt (Abb. 4).

Abbildung 4: Drei Schweregrade für Usability-Probleme



Die drei Stufen wurden in Anlehnung an Barnum (2011: 264) erstellt. Die einzelnen Symbole haben folgende Bedeutungen:

- Rot = Schwerwiegendes Problem:
 - Der Benutzer kann die Aufgabe nicht beenden.
 - Der Benutzer kann die Aufgabe nur mit sehr viel Mühe beenden und ist extrem verwirrt oder braucht sogar Hilfe, um die Aufgabe zu lösen.
- Orange = Wesentliches Problem:
 - Der Benutzer kann die Aufgabe zwar lösen, ist aber sehr frustriert.
 - Der Benutzer ist sehr unzufrieden und drückt seinen Unmut aus oder äussert einen Wunsch.
- Gelb = Marginales Problem:
 - Benutzer zögern oder wählen zuerst einen falschen Weg, aber bemerken den Fehler und können ihn korrigieren.
 - Der Benutzer ist leicht irritiert oder leicht verärgert, kann das Problem aber ohne grössere Schwierigkeiten lösen.

Die Usability-Probleme werden beschrieben und Verbesserungsvorschläge sollen dabei helfen, mögliche Lösungen aufzuzeigen. In blauen Kästen werden Beispiele für das Problem gezeigt, um es zu verbildlichen. Abgesehen von den Symbolen für den Schweregrad des Problems stammen alle Abbildungen von der Webseite Swissbib. Der Anhang H bietet eine zusätzliche Übersicht und zeigt, mit welchem der beiden Usability-Tests das Problem identifiziert wurde. Die Nutzungsprobleme sind im folgenden Kapitel nach Thematik geordnet und im Anhang H werden sie in der Reihenfolge des Schweregrades aufgelistet.

5.1 Menüleiste

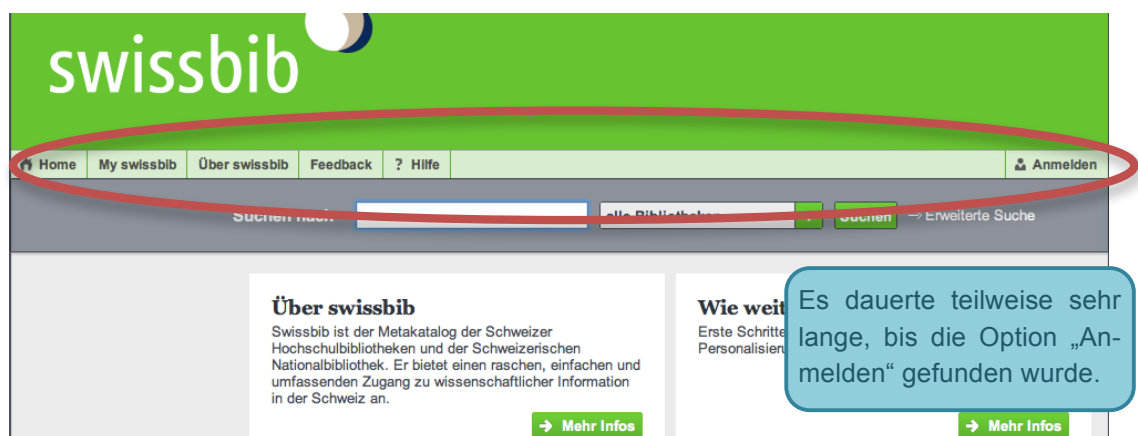
Nr.1: Grösse der Menüleiste



Problem: Die Benutzer haben Mühe sich in der Menüleiste zurechtzufinden und erkennen die unterschiedlichen Optionen schlecht (Abb. 5).

Vorschlag: Die gesamte Menüleiste sollte verbreitert werden. Die Symbole und die Schrift sollten vergrößert werden, damit die unterschiedlichen Menüs schneller und besser wahrgenommen werden.

Abbildung 5: Menüleiste Swissbib



5.2 Autovervollständigung

Nr. 2: Autovervollständigung verschwindet



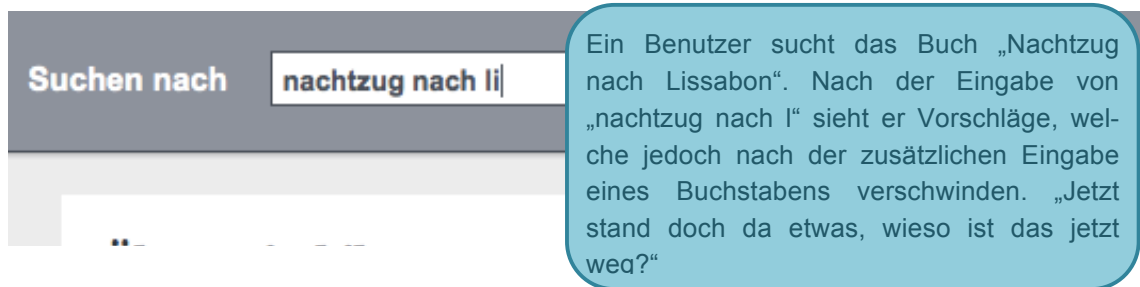
Problem: Bei der Eingabe von Text im Suchfeld erscheinen Vorschläge. Diese Autovervollständigung wird nicht immer angezeigt und verschwindet bei gewissen Anzahlen Buchstaben. Dies ist irritierend für die Benutzer (Abb. 6 und 7).

Vorschlag: Vorschläge sollten, auch wenn Buchstaben hinzugefügt werden, immer angezeigt werden.

Abbildung 6: Autovervollständigung „nachtzug nach l“



Abbildung 7: Fehlende Autovervollständigung „nachtzug nach li“



Nr. 3: Übernahme der Autovervollständigung



Problem: Wenn einer der Vorschläge der Autovervollständigung angewählt wird, wird nicht immer der gesamte Ausdruck in das Suchfeld übernommen. Es wird nur nach einigen Wörtern gesucht und nicht nach dem gesamten Ausdruck. Das führt zu unpräzisen Suchergebnissen (siehe Abb. 8 und 9).

Vorschlag: Der gesamte gewählte Ausdruck sollte in das Suchfeld übernommen werden, damit die Suchresultate möglichst genau dem Vorschlag entsprechen.

Abbildung 8: Autovervollständigung "du sollst"

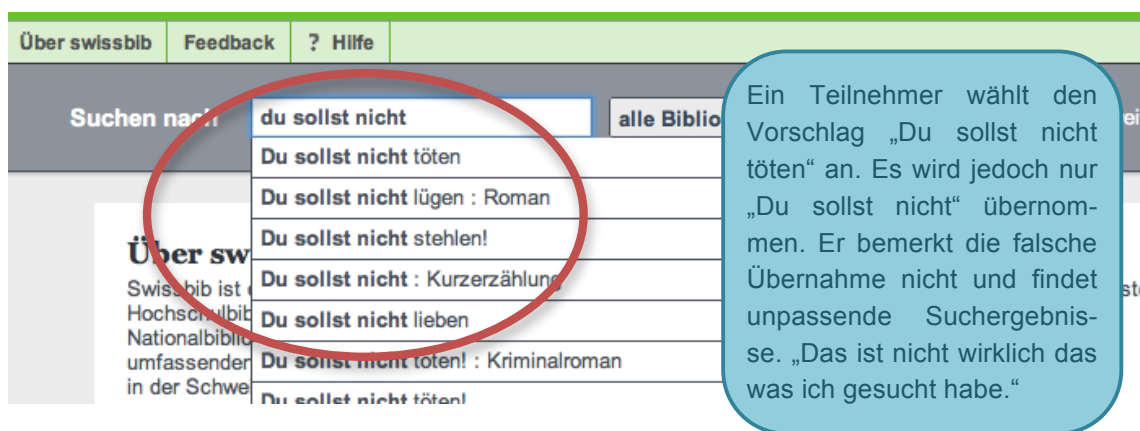
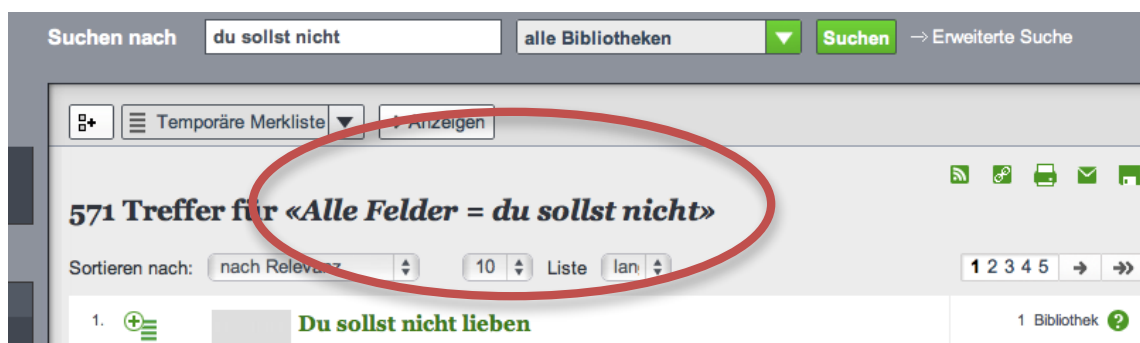


Abbildung 9: Übernahme der Autovervollständigung „du sollst nicht töten“



5.3 Null-Treffermengen

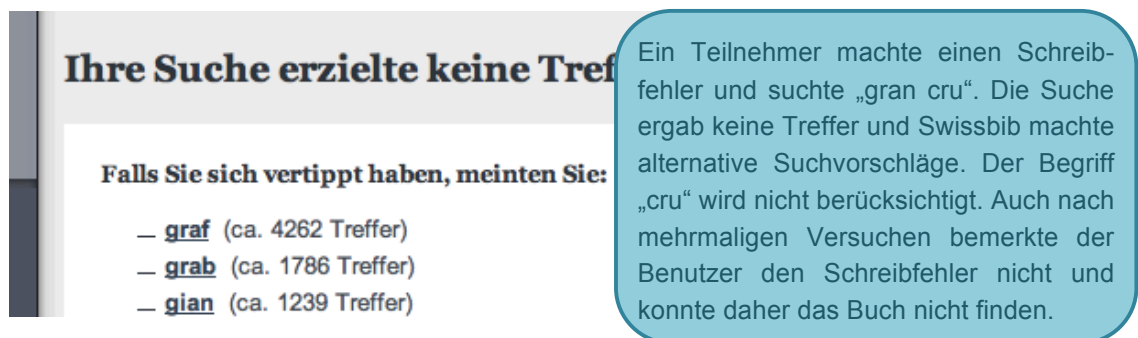
Nr. 4: Unrelevante Vorschläge bei Null-Treffermengen



Problem: Wenn Benutzer Schreibfehler bei ihren Suchanfragen machen und dadurch keine Resultate erhalten, liefert Swissbib alternative Vorschläge. Jedoch entsprechen diese Vorschläge häufig nicht dem, was der Benutzer gesucht hat (Abb. 10).

Vorschlag: Die Datenverarbeitung sollte verbessert werden, um dem Benutzer besonders ähnliche Alternativen zu liefern. Es sollten dabei alle eingegebenen Begriffe berücksichtigt werden.

Abbildung 10: Alternative Suchvorschläge „gran cru“



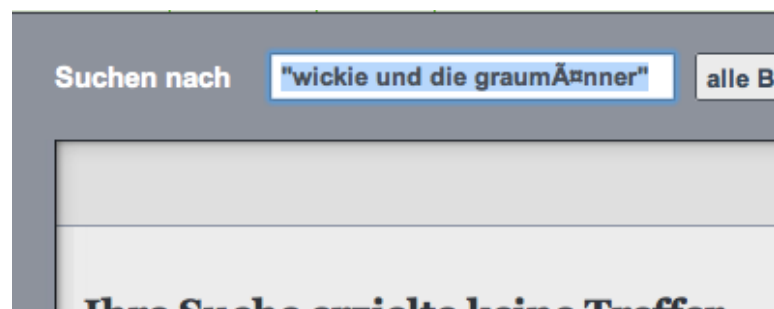
Nr. 5: Verarbeitung von Umlauten



Problem: Beim Remote Test hatten einige Teilnehmer Probleme mit der Eingabe von Umlauten. Es handelt sich wahrscheinlich um ein Problem mit der Zeichenkodierung (Abb. 11). Möglich wäre, dass diese Schwierigkeit durch die Testumgebung generiert wurde. Der genaue Grund konnte jedoch nicht ermittelt werden. Ein Problem beim Kopieren und Einfügen des Textes konnte ausgeschlossen werden, da sich teilweise auch Schreibfehler in den Eingaben fanden.

Vorschlag: Die Ursache für das Problem identifizieren und beheben.

Abbildung 11: Umlautverarbeitung auf Loop11



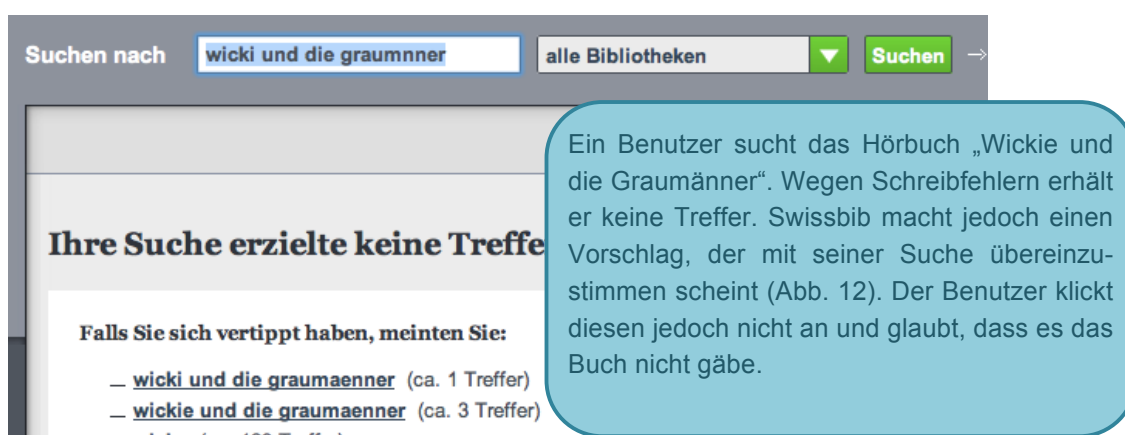
Nr. 6: Vorschläge bei Null-Treffermengen zu klein



Problem: Falls eine Suche keine Resultate liefert bietet Swissbib Vorschläge an. Auch wenn diese scheinbar der eigentlichen Suche entsprechen, werden sie nicht genutzt.

Vorschlag: Die Aufmerksamkeit der Nutzer sollte mehr auf die Alternativen gelenkt werden. Der Text „Falls Sie sich vertippt haben, meinten Sie“ sollte grösser geschrieben werden als die eher negative Aussage „Ihre Suche erzielte keine Treffer“. Auch für die Vorschläge könnte eine grössere Schrift gewählt werden.

Abbildung 12: Alternative Suchvorschläge „wicki und die graummer“



5.4 Suchresultate

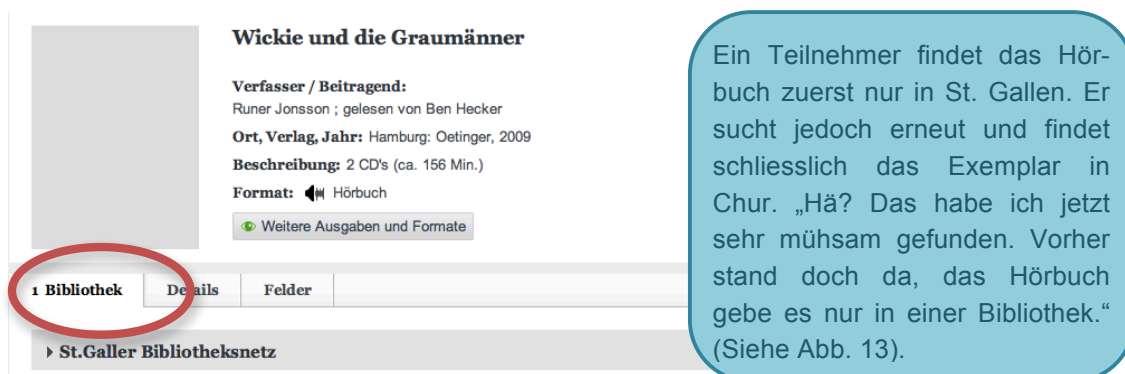
Nr. 7: Dublettenbereinigung



Problem: Da nicht alle Exemplare eines Dokuments zu einem Datensatz verbunden sind, können Benutzer nicht alle Bibliotheken, die ein Dokument besitzen, identifizieren.

Vorschlag: Die Dublettenbereinigung sollte noch verbessert werden.

Abbildung 13: Datensatz „Wickie und die Graumänner“ für St. Gallen



Nr. 8: Sortierung nach Relevanz



Problem: In der Trefferliste werden die relevanten Dokumente nicht immer am Anfang angezeigt, dadurch finden Benutzer das gesuchte Dokument weniger schnell (Abb. 14).

Vorschlag: Bei der Sortierung nach Relevanz sollten Dokumente, bei denen die Suchanfrage genau dem Titel entspricht, höher gerankt werden.

Abbildung 14: Erste Suchresultate "grand cru"



Ein Teilnehmer sucht das Buch mit dem Titel „Grand cru“. Die ersten Ergebnisse haben keinen der beiden Begriffe im Titel. „Oh, da bekomme ich sehr viele Resultate. Aber das ist nicht, was ich gesucht habe“.

Ein Online Teilnehmer kommentiert: „Das Buch war überraschend weit unten auf der Trefferliste.“

Nr. 9: Unauffällige Fassettennavigation



Problem: Die Fassettennavigation auf der linken Seite wurde schlecht wahrgenommen (Abb. 15). Dadurch wurde sie selten genutzt und Benutzer griffen eher auf die erweiterte Suche zurück, um ihre Resultate einzugrenzen.

Vorschlag: Die Fassettennavigation sollte sich durch andere, hellere Farben besser vom dunklen Hintergrund abheben. Zusätzlich könnten die Titel der einzelnen Fassetten (Verbund, Verfasser etc.) grösser geschrieben werden.

Abbildung 15: Fassettennavigation



Ein Teilnehmer sucht nach deutschen Dokumenten mit dem Begriff „Bioenergie“. Er sieht aber Fassetten nicht und meint: „Ich fände es noch gut, wenn man irgendwo Dinge wählen könnte.“ Die Fassetten findet der Teilnehmer erst nach der Konsultation der Hilfe.

5.5 Trefferliste

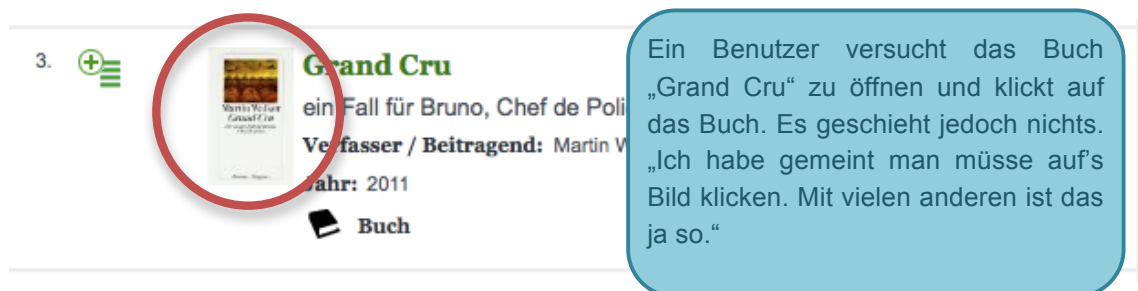
Nr. 10: Öffnen des Dokuments mit Bild



Problem: Um ein Dokument zu öffnen, klickten mehrere Teilnehmer das Bild des Umschlags an (Abb. 16). Dieses hat jedoch keine Funktion.

Vorschlag: Es wäre sinnvoll, wenn das Umschlagsbild ebenfalls dazu dienen würde, ein Dokument zu öffnen. Dies erleichtert die Navigation zu den Details eines Dokuments.

Abbildung 16: Umschlagbild „Grand Cru“



Nr. 11: Bibliothek identifizieren



Problem: Auf der Trefferliste besteht die Möglichkeit abzufragen, welche Bibliotheken ein Dokument besitzen (siehe Abb. 17). Jedoch wird diese Information von den meisten Nutzern nicht wahrgenommen.

Vorschlag: Das Symbol könnte vergrößert werden. Der Text könnte ersetzt werden durch „In diesen x Bibliotheken:“

Abbildung 17: Anzeige der Bibliotheken auf der Trefferliste



Nr. 12: Unklares Merksymbol



Problem: Das Merksymbol, welches eigentlich dazu dient, ein Dokument zur Merkliste hinzuzufügen, wird falsch interpretiert oder nicht wahrgenommen. Nutzer versuchen damit, das Dokument zu öffnen oder haben das Gefühl, es würde zum Vergrößern dienen (Abb. 18).

Vorschlag: Die Funktion des Symbols sollte klarer gemacht werden zum Beispiel durch einen Verweis auf die Funktion beim Darüberfahren (wie zum Löschen, siehe Abb. 19). Besser wäre es, im Symbol den Text „Merken“ zu integrieren.

Abbildung 18: Symbol zum Hinzufügen eines Dokuments auf die Merkliste

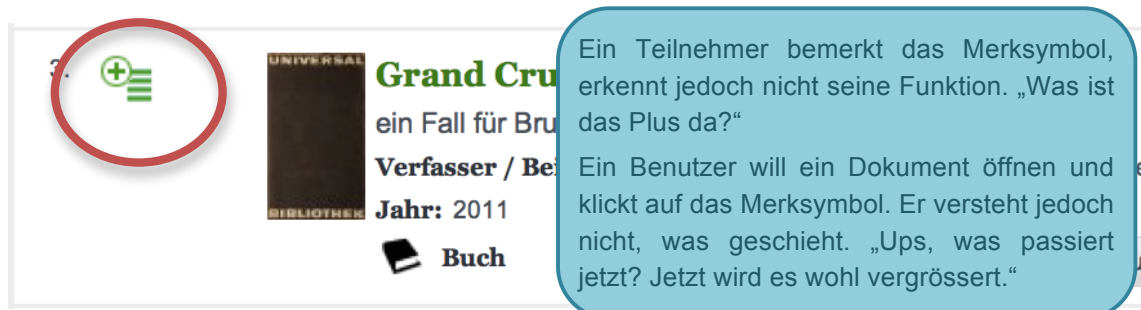


Abbildung 19: Symbol zum Löschen eines Dokuments aus der Merkliste



5.6 Erweiterte Suche

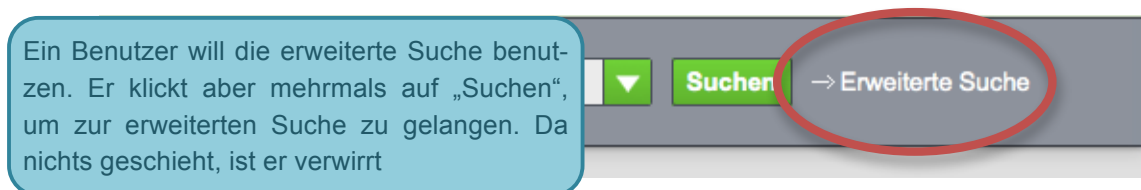
Nr. 13: Unklarer Link zur erweiterten Suche



Problem: Der Link zur erweiterten Suche ist nicht für jeden Benutzer als solcher erkennbar. Zudem verleitet der Pfeil neben dem Suchen-Button zu der Annahme, dass beim Wählen des Suchen-Buttons die erweiterte Suche aktiviert wird (siehe Abb. 20).

Vorschlag: Der Text „Erweitere Suche“ könnte unterstrichen werden oder die Farbe beim Darüberfahren könnte sich ändern.

Abbildung 20: Link zur erweiterten Suche



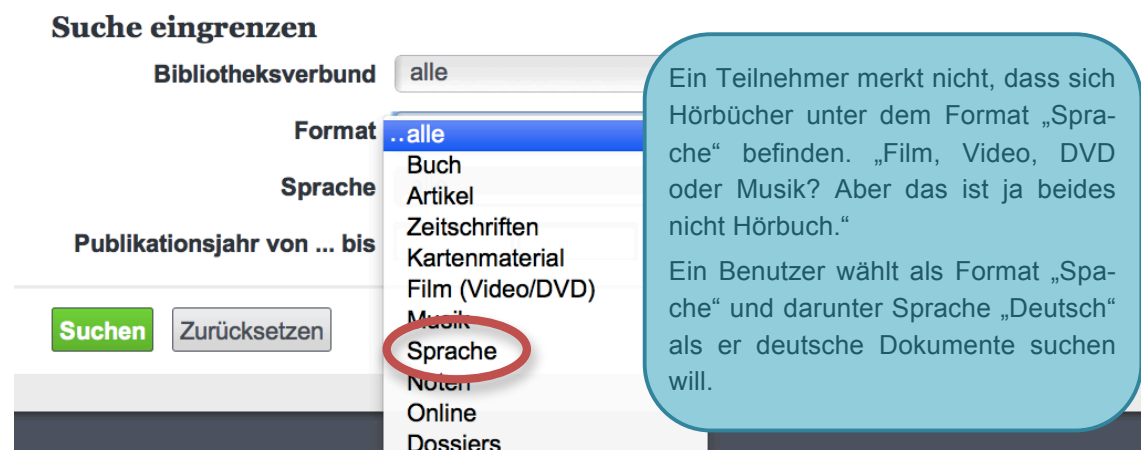
Nr. 14: Bezeichnung des Formats „Sprache“



Problem: Bei der erweiterten Suche ist das Format mit der Bezeichnung „Sprache“ unverständlich und wird deshalb nicht oder falsch genutzt (Abb. 21).

Vorschlag: Die Bezeichnung „Sprache“ sollte geändert werden. Eine Alternative wäre zum Beispiel „Ton / Hörbuch“.

Abbildung 21: Auswahl der Formate in der erweiterten Suche



5.7 Bestandesinformationen

Nr. 15: „Zum Bestand“ und „Informationen zum Bestand“



Problem: Der Link „zum Bestand“ öffnet die Webseite der Bibliothek, die das Dokument besitzt und die Option „Informationen zum Bestand“ zeigt die Signatur und Verfügbarkeit des Dokuments an (Abb. 22). Jedoch können die Benutzer die Funktionen nicht aus den Begriffen ableiten und wissen nicht, wo sie klicken sollen. Viele Benutzer klicken auf „zum Bestand“ und sehen die von Swissbib gelieferten Informationen zur Verfügbarkeit oder Signatur nicht.

Vorschlag: Es wäre sinnvoll, wenn der Button „Informationen zum Bestand“ gelöscht würde und Informationen über die Verfügbarkeit oder die Signatur direkt angezeigt würden (Abb. 23). Dadurch muss der Nutzer zudem einmal weniger klicken, um zu den Informationen zu gelangen.

Abbildung 22: Optionen für nähere Informationen über ein Dokument

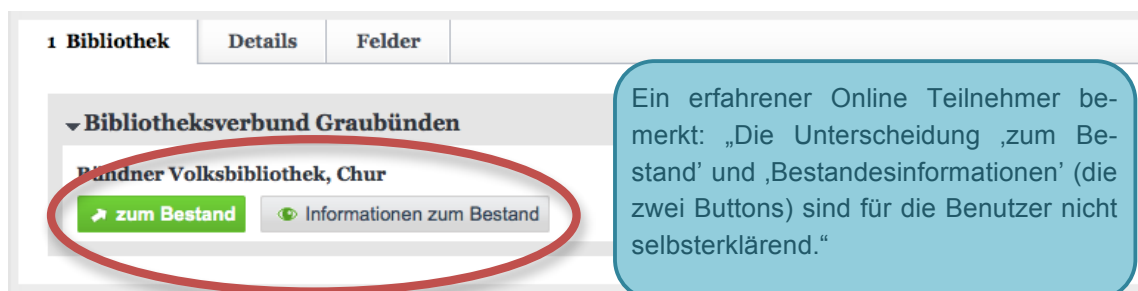


Abbildung 23: Informationen zum Churer Hörbuch „Wickie und die Graumänner“



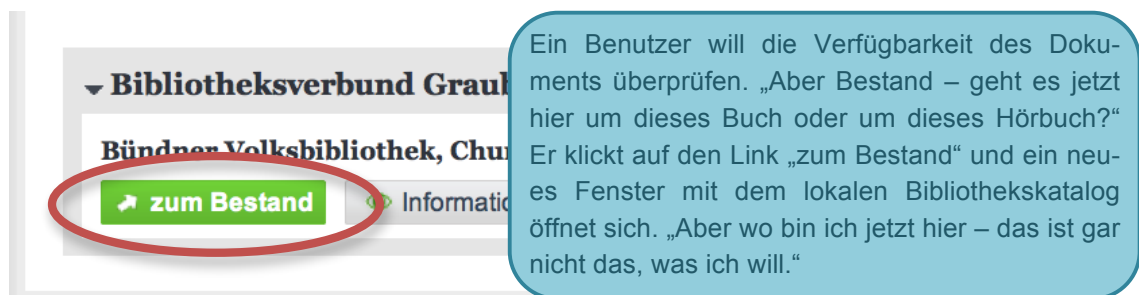
Nr. 16: Bezeichnung „Zum Bestand“



Problem: Der Begriff „Bestand“ ist für einen Nicht-Bibliothekar nicht aussagekräftig (Abb. 24). Es ist nicht klar, dass der Link „Zum Bestand“ zu einer externen Bibliothek führt und ungewohnte Nutzer sind verwirrt, wenn sich ein neues Fenster öffnet.

Vorschlag: Der Link könnte umbenannt werden in „zur Bibliothek“ oder „zur Bibliothek, die das Dokument besitzt“.

Abbildung 24: Button „zum Bestand“



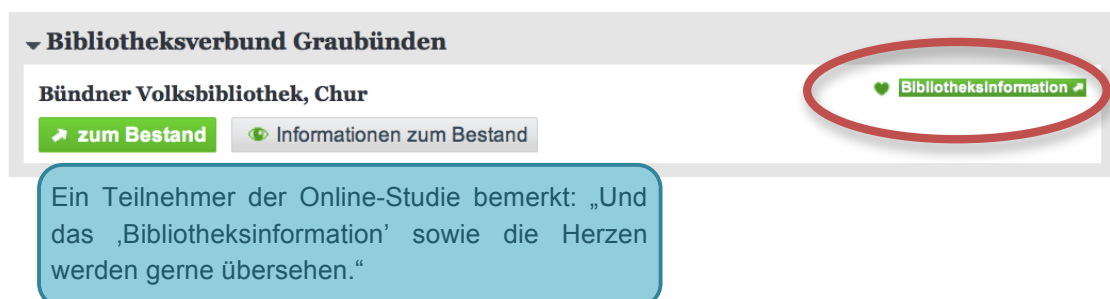
Nr. 17: Informationen über die Bibliothek



Problem: Auf der Aufnahme eines Exemplars bestehen die Möglichkeiten, die besitzende Bibliothek als Favorit zu definieren (Herzsymbol) und Informationen über diese Institution zu erhalten. Jedoch sind die Symbole sehr unauffällig und der Text kann nur schwer gelesen werden (Abb. 25).

Vorschlag: Beide Buttons sollten vergrößert werden, damit sie mehr auffallen und besser gelesen werden können.

Abbildung 25: Buttons „Bibliotheksinformation“ und „Favorit-Bibliothek“



5.8 Anzeige mit Internet Explorer 8

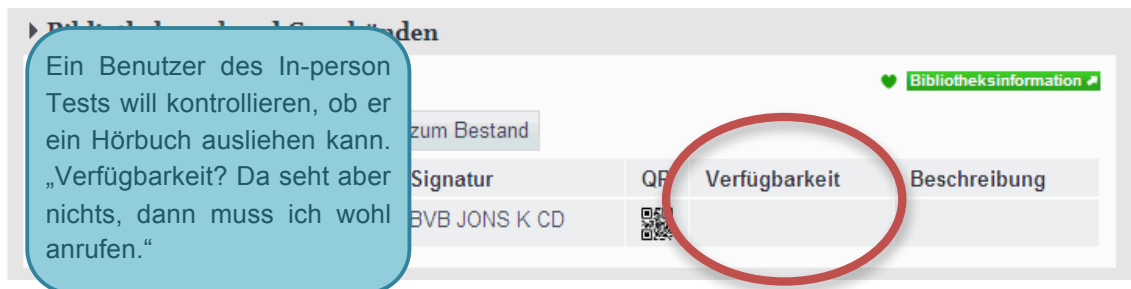
Nr. 18: Anzeige der Verfügbarkeit mit dem Internet Explorer 8



Problem: Der Verfügbarkeitsstatus auf Swissbib wird bei der Verwendung des Internet Explorers 8 nicht angezeigt (siehe Abb. 26).

Vorschlag: Swissbib sollte für den Internet Explorers 8 optimiert werden.

Abbildung 26: Anzeige der Verfügbarkeit mit dem Internet Explorer 8



Nr. 19: Anzeige des Marc-Formats mit dem Internet Explorer 8



Problem: Bei der Benutzung des Internet Explorers 8 werden die drei Reiter „Bibliotheken“, „Details“ und „Felder“ nicht wie üblich nebeneinander (Abb. 27) sondern untereinander angezeigt (Abb. 28). Die Anzeige des Marc-Formats ist für einen Nicht-Bibliothekaren verwirrend.

Vorschlag: Swissbib sollte für den Internet Explorers 8 optimiert werden.

Abbildung 27: Anzeige des Marc-Formats mit dem Internet Explorer 8

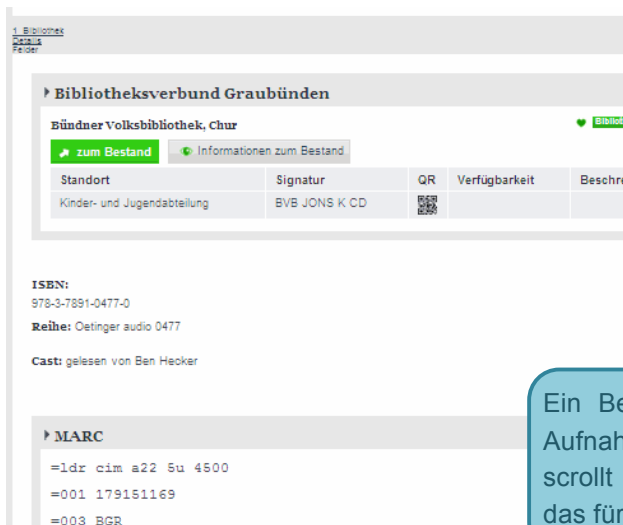
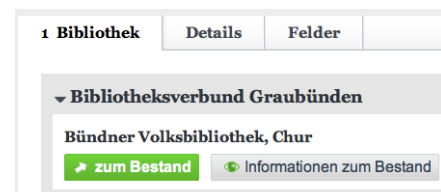


Abbildung 28: Richtige Anzeige des Marc-Formats mit Firefox



Ein Benutzer befindet sich auf der Aufnahme eines Dokuments und scrollt nach unten. „Was sind denn das für komische Zeichen?“

5.9 Hilfe-Themen

Nr. 20: Organisation der Hilfe-Themen



Problem: Der Unterschied zwischen den Seiten „Wie weiter?“, „Über swissbib“ und „Hilfe“ ist nicht klar (Abb. 29). Viele Benutzer suchen unter einer Option und wenn sie die Information dort nicht finden, geben sie auf und suchen nicht noch an einer zweiten Stelle.

Vorschlag: Die beiden Menüoptionen „Wie weiter?“ und „Hilfe“ sollten in einer Seite vereinigt werden. Zusätzlich könnten die grossen Felder in der Mitte der Seite umbenannt werden, damit deren Inhalt klarer wird: „Was ist swissbib?“ und „Wie funktioniert swissbib?“.

Abbildung 29: Startseite mit unterschiedlichen Hilfeseiten



Ein Teilnehmer macht folgenden Kommentar: „Die ‚Hilfe‘, bzw. die FAQ fand ich nicht gut organisiert. Wenn ich wissen möchte, was die Vorteile eines Kontos sind, suche ich eher nicht unter ‚Hilfe‘, sondern unter ‚Über uns‘ oder so.“

Nr. 21: Navigation innerhalb der „Hilfe“



Problem: Benutzer haben Mühe sich in der umfangreichen Hilfe zurechtzufinden und können die Hilfethemen schlecht lesen (Abb. 30).

Vorschlag: Für die Anzeige der Hilfethemen sollte eine grössere Schrift gewählt werden. Denkbar wäre auch, die Hilfethemen auf der linken Seite als Navigationsleiste anzufügen. Die Navigationsleiste müsste dabei aber grösser und heller erscheinen (vgl. Kapitel 5.4, Problem Nr. 9).

Abbildung 30: Navigation innerhalb der Hilfe



Ein Teilnehmer sucht in der Hilfe Informationen zum Filtern von Dokumenten und hat Mühe sich zurechtzufinden. „Moment – wo habe ich das jetzt vorher gesehen, war das hier?“

Nr. 22: Lesbarkeit im Menü „Hilfe“



Problem: Die Hilfe ist zu textlastig (siehe Abb. 31) und Benutzer haben keine Lust sich näher mit einem Thema zu beschäftigen.

Vorschlag: Die Texte sollten vereinfacht und gekürzt werden. Ausserdem würde die Wahl einer grösseren Schrift mit mehr Leerraum („white space“) das Lesen am Bildschirm vereinfachen.

Abbildung 31: Hilfeseite

Ein Benutzer befindet sich auf der Hilfeseite und meint dazu: „Oh Gott, das ist ja sehr kompliziert. Und dann ist das noch so klein, ich muss mich richtig anstrengen, das zu lesen.“

Ergebnisse sortieren und filtern

Die Ergebnisse jeder Suche erscheinen als **Trefferliste**. Die **Reihenfolge** der Treffer können Sie mit dem Menü **Sortieren nach** beeinflussen, nämlich nach Relevanz, Jahr, Autorennamen oder Titel. Falls Sie mehr Treffer auf einer Seite angezeigt haben möchten, können Sie die maximale Anzahl der Treffer und das Format der Liste (lang/kurz) verändern.

Auf der linken Seite der Trefferliste stehen Ihnen **Filter** zur Verfügung, mit denen Sie Ihre Ergebnisse **beliebig eingrenzen** können. Neben den schon bei der Suche verfügbaren Eingrenzungskriterien (Bibliotheksverbund, Autor, Format, Sprache, Jahr), haben Sie hier zudem die Möglichkeit, den **Inhalt** der gesuchten Medien feiner nach einer bestimmten Person, einem Ort, einer Publikationsform oder einer bestimmten Zeitspanne einzugrenzen.

Wenn Sie **mehrere Filter** anklicken, bildet sich aus diesen eine **Schnittmenge**. Filter können jederzeit im oberen Teil der Spalte anhand des Löschsymbols wieder einzeln entfernt werden.

Bei gewissen Titeln erscheint ein Link zu **Weitere Ausgaben und Formaten**. Dieser bietet Ihnen eine Liste von Titeln, die **Ähnlichkeiten** mit dem angeklickten Dokument aufweisen, sei es z.B., weil es sich um verschiedene Auflagen eines Werkes handelt oder weil ein Dokument in mehreren Formaten zugänglich ist (Druckschrift, elektronische Ressource, Film, Ton etc.). Dies ist dann hilfreich, wenn zum Beispiel die gewünschte Ausgabe in Ihrer Bibliothek nicht vorhanden oder bereits ausgeliehen ist.

Achtung: Unter den weiteren Ausgaben und Formaten können sich Dokumente befinden, die mit dem ausgewählten Titel eigentlich identisch sind. Das Zusammenführen von identischen Titeln funktioniert in einigen Fällen noch nicht zufriedenstellend, vor allem bei Zeitschriften: [Beispiel hier in neuem Fenster](#).

Beispiel:
Suche nach »Biochemistry«, dann Eingrenzung der Trefferliste anhand der Suchfilter **Format** (»Tagungsband«), **Jahr** (seit 2005) und **inhaltliches Thema** (»Immunology«).

Beispiel:
Suche nach »Frisch« in der einfachen Suche. Der **Autor-Filter** erlaubt Ihnen, alle von Max Frisch geschriebenen Titel anzuzeigen. Der **Inhalt-Filter** (Person) ergibt eine Trefferliste von Dokumenten über Max Frisch.

5.10 My swissbib

Nr. 23: Unterschied „My swissbib“ und „Anmelden“



Problem: Der Unterschied zwischen den beiden Menüoptionen „My swissbib“ und „Anmelden“ wird nicht oder nur schlecht verstanden. Intuitiv vermuten die Benutzer die Anmeldeoption unter dem Menü „My swissbib“. Hier finden sie aber kein passendes Untermenü (Abb. 32).

Vorschlag: Alle personalisierbaren Menüoptionen sollten sich an derselben Stelle befinden. So könnte der Reiter „My swissbib“ links neben der Option „Anmelden“ platziert werden.

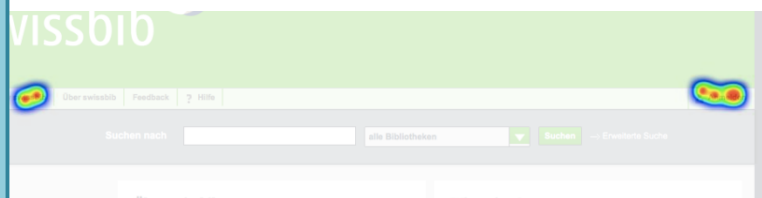
Abbildung 32: Menüoptionen „My swissbib“ und „Anmelden“



Teilnehmer will sich anmelden. „My swissbib, das müsste ja das sein, ich kann zwar kein Englisch.“
Öffnet Menü. „Nein, das kann es nicht sein.“

Auch die von Loop11 generierte Heatmap zeigt diese Verwirrung (Abb. 33).

Abbildung 33: Heatmap "My swissbib" und "Anmelden"



Nr. 24: Anmeldeverfahren



Problem: Nicht-Studenten sind verwirrt über den Term „Switch aai“ bei den Anmeldeverfahren (Abb. 34). Teilweise klicken die Benutzer direkt auf „Anmelden“ und versuchen sich über „Switch aai“ anzumelden, was aber nicht funktioniert, da sie dort nicht eingeschrieben sind.

Vorschlag: Das Anmeldeverfahren „Switch aai“ sollte an zweiter Stelle stehen. Die Bezeichnungen könnten geändert werden in:

- Normale Anmeldung
- StudentInnen und Universitätsangehörige (via Switch aai)

Abbildung 34: Loginvarianten

Ein Teilnehmer überlegt, welches Verfahren er wählen soll. „Äh, muss ich jetzt das ändern?“

Einige Teilnehmer haben es gar nicht geschafft sich anzumelden. „Es ist mir nicht gelungen mich einzuloggen mit den Angaben, die Sie geliefert haben.“

Nr. 25: Zugriff auf das Benutzerkonto



Problem: Einmal angemeldet, hat der Benutzer die Möglichkeit auf sein Konto via Nutzernamen oder via Reiter „My swissbib“ zuzugreifen. Jedoch sind sich die Benutzer unsicher, wie sie auf ihr Konto gelangen können.

Vorschlag: Statt „Nutzername: E-Mail“ könnte neben dem Symbol stehen „My swissbib: E-Mail“. Der Reiter „My swissbib“ wird überflüssig, sobald ein Benutzer angemeldet ist.

Abbildung 35: Zugriff auf das Benutzerkonto

Ein Teilnehmer sieht die E-Mailadresse. „Soll ich mal da klicken, ich weiss zwar nicht, was dann passiert.“

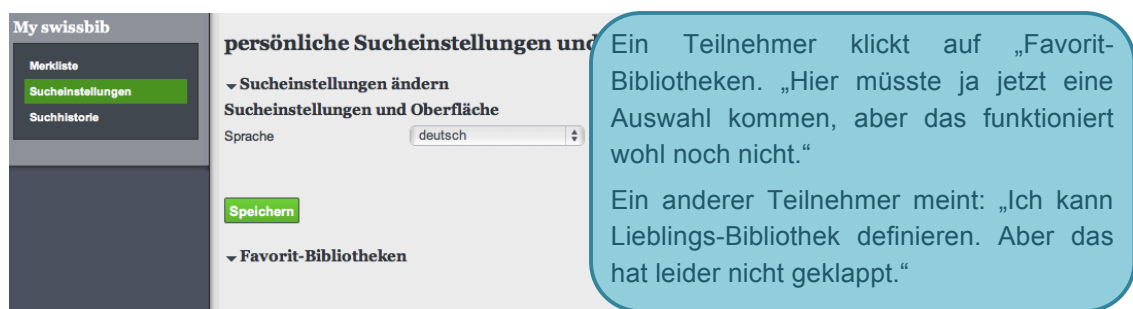
Nr. 26: Festlegen der Favorit-Bibliotheken



Problem: Unter den Sucheinstellungen können keine Favorit-Bibliotheken definiert werden (siehe Abb. 36). Diese Funktion ist einzig zugänglich durch ein Untermenü von „My swissbib“ und wird so nicht oder nur schwer gefunden.

Vorschlag: Favoriten sollten unter den Sucheinstellungen definiert werden können. Falls dies nicht möglich ist, sollte die Anzeige „Favorit-Bibliothek“ aus den Sucheinstellungen entfernt und in das seitliche Menü links integriert werden.

Abbildung 36: Favorit-Bibliotheken auf den persönlichen Sucheinstellungen



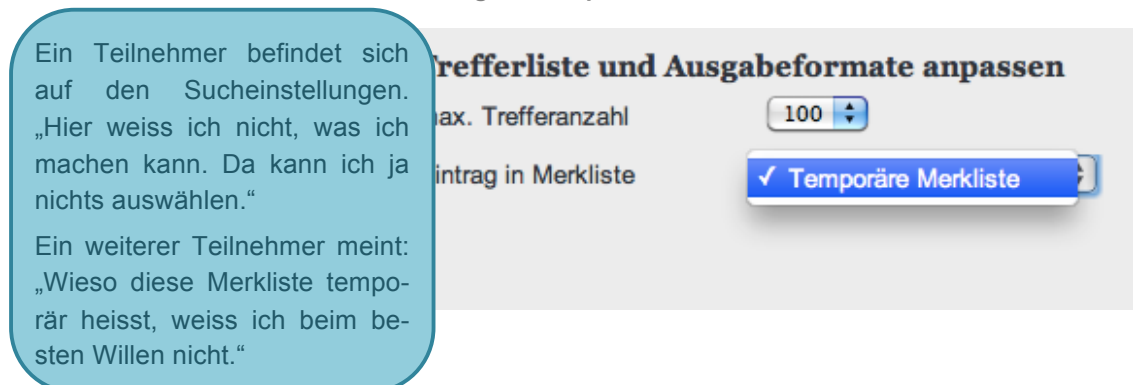
Nr. 27: Bezeichnung „Temporäre Merkliste“



Problem: Die Bezeichnung „Temporäre Merkliste“ ist für die Benutzer unklar und die Bedeutung wird nicht verstanden (Abb. 37).

Vorschlag: Der Begriff „temporär“ sollte weggelassen werden. Ein zusätzlicher Text könnte die Funktion genauer erklären: „Die Dokumente in dieser Liste werden nur für die laufende Sitzung gespeichert. Wenn Sie Dokumente für längere Zeit speichern möchten, melden Sie sich an und erstellen Sie eine neue Merkliste.“

Abbildung 37: Temporäre Merkliste



6 Resultate aus dem Methodenvergleich

Die gesammelten Daten der zwei Usability-Tests vermögen nicht, Aussagen über die gesamte Nutzerschaft von Swissbib zu treffen. Auch wenn die Auswahl der Teilnehmer beim In-person Usability-Test verschiedene Nutzertypen abdeckte, war sie mit fünf Personen sehr klein. Zudem erreichte die Rekrutierungsmethode des Remote Tests vor allem Spezialisten aus dem Bereich Information und Dokumentation und deckte damit andere Nutzergruppen wie zum Beispiel Studenten nicht ab. Diese Auswahlverfahren werden aber gerade durch die jeweilige Usability-Methode bedingt. Ein Discount Usability-Test findet nur mit wenigen Personen statt und ein Remote Test bedient sich oft Datenbanken oder Verteilerdiensten, um Personen zu rekrutieren. Diese haben selten genau dieselbe Verteilung der Merkmale der Gesamtnutzerschaft eines Produkts.

Die hier dargestellten Resultate erheben damit keinen Anspruch auf Repräsentativität. Damit statistisch repräsentative Aussagen getroffen werden könnten, müssten die Tests mit grösseren, repräsentativeren Stichproben wiederholt werden. Dies würde den Rahmen dieser Arbeit jedoch sprengen und nicht mit den Testmethoden übereinstimmen. Die Datenanalyse soll zeigen, wo Unterschiede zwischen den beiden untersuchten Methoden liegen und was die möglichen Ursachen für diese Unterschiede sein könnten.²⁶

Um die Unterschiede in der Grösse der Stichprobe und die damit verbundene Unsicherheit der Resultate darzustellen, werden Diagramme mit Konfidenzintervallen repräsentiert. Je grösser eine Stichprobe, desto kleiner das Konfidenzintervall oder anders gesagt: je mehr Personen befragt wurden, desto grösser ist die Vertrauenswürdigkeit der Daten (Albert, Tullis, Tedesco, 2010: 128). Wie von Albert, Tullis und Tedesco (2010: 129) vorgeschlagen, wurde ein Alpha Level²⁷ von 90% gewählt. Berechnet wurden die Konfidenzintervalle mit den von Jeff Sauro zur Verfügung gestellten Online Tools.^{28,29,30}

²⁶ WIKIA. Repräsentativität. In: *Marktforschungs-Wiki* [online]. <http://marktforschung.wikia.com/wiki/Repräsentativität> (letzter Zugriff: 26.06.2013)

²⁷ Dieser Wert legt fest, wie hoch die Irrtumswahrscheinlichkeit sein darf.

²⁸ Für kontinuierliche Daten: SAURO, Jeff. Confidence interval calculator. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 12.03.2012. <http://www.measuringusability.com/ci-calc.php&summary=1> (letzter Zugriff: 26.06.2013)

²⁹ Für Erfolgsraten: SAURO, Jeff. Confidence interval calculator for a completion rate. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 01.10.2005. <http://www.measuringusability.com/wald> (letzter Zugriff: 26.06.2013)

³⁰ Für Aufgabenzeiten: SAURO, Jeff. Graph and calculator for confidence intervals for task times. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 06.02.2006. http://www.measuringusability.com/time_intervals (letzter Zugriff: 26.06.2013)

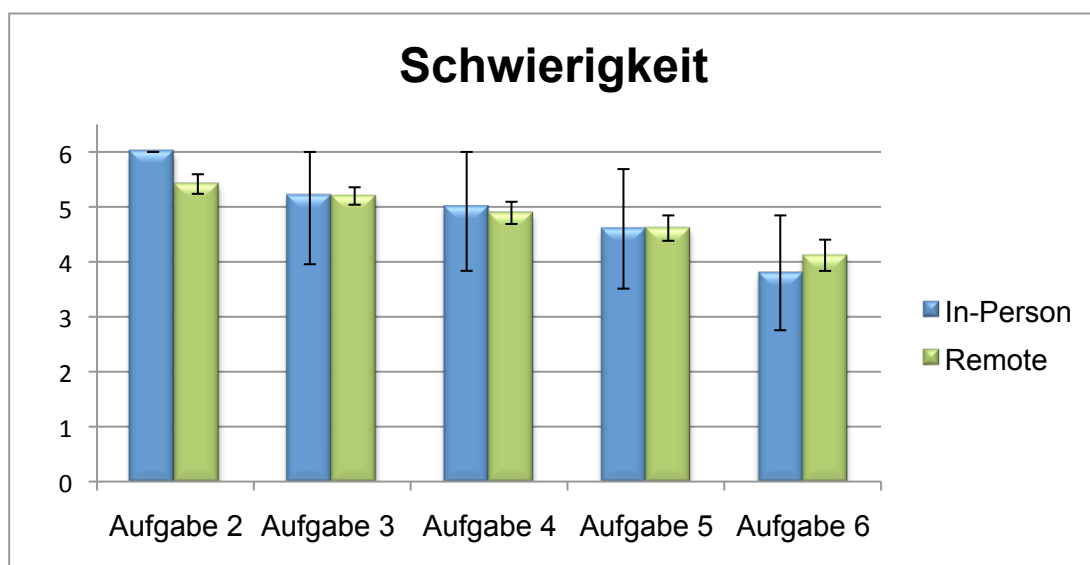
6.1 Überprüfung der subjektiven Einstellung

Die Hypothese 1 lautete wie folgt: „Die subjektive Einstellung der Testpersonen während eines unmoderierten Remote Usability-Tests ist nicht identisch mit der Einstellung während eines In-person Usability-Test, da die Umgebungen die Testpersonen unterschiedlich beeinflussen.“ Die Validität dieser Aussage soll durch die Single Ease Question am Ende jeder Aufgabe, durch die System Usability Scale (SUS) und durch einen zweiten Posttest-Fragebogen geprüft werden.

6.1.1 Single Ease Question

Mit der Single Ease Question (SEQ) wurde gemessen, wie schwierig die einzelnen Teilnehmer die unterschiedlichen Aufgaben einschätzten. Die Tabelle 6 zeigt diese Einschätzung für die einzelnen Aufgaben. Null bedeutet die Aufgabe wurde sehr schwierig empfunden und sechs sehr einfach.

Tabelle 6: Einschätzung des Schwierigkeitsgrads

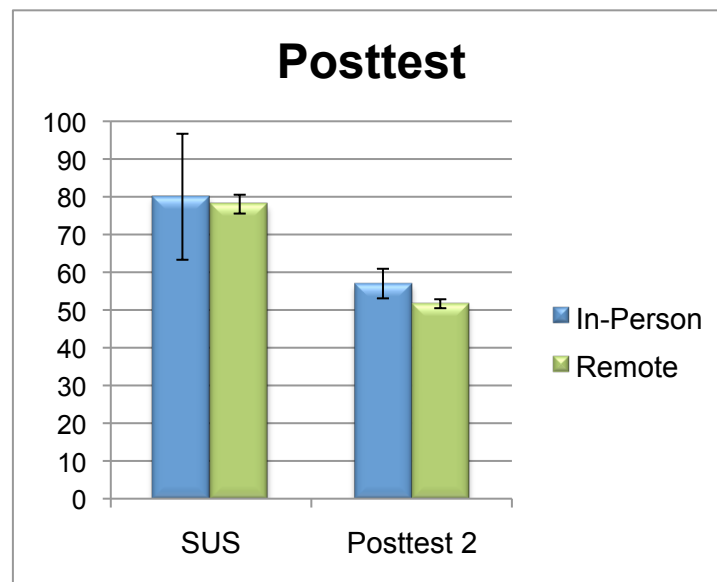


Bei der ersten Aufgabe wurde nicht nach der Schwierigkeit gefragt, da diese dazu diente, einen ersten Eindruck zu erfassen. Entsprechend dem Aufbau des Tests wird jede weitere Aufgabe ein wenig schwieriger eingeschätzt. Trotz der kleinen Datenmenge aus dem In-person Test liegen die Mittelwerte erstaunlich nahe beieinander und es gibt keine signifikanten Unterschiede. Es gibt eine kleine Tendenz, dass die Teilnehmer beim In-person Nutzertest die einfachste Aufgabe einfacher und die schwierigste schwieriger einschätzen.

6.1.2 System Usability Scale

Die Tabelle 7 zeigt die Resultate der beiden Posttest-Fragebögen. Bei der SUS konnten maximal 100 Punkte für die allgemeine Benutzerfreundlichkeit erreicht werden. Für den zweiten Posttest-Fragebogen, der die Einstellung der Teilnehmer gegenüber der Testmethode mass, konnten bei positiver Bewertung maximal 60 Punkte erreicht werden. Benutzerfreundlichkeit als auch Testmethode wurden in beiden Tests sehr gut bewertet. Der SUS-Wert ist, verglichen mit anderen Webseiten, sogar überdurchschnittlich hoch (Albert, Tullis, Tedesco, 2010: 152). Erläuterungen zu diesem guten Abschneiden finden sich im Kapitel 5.

Tabelle 7: Resultate der Posttest-Fragebögen

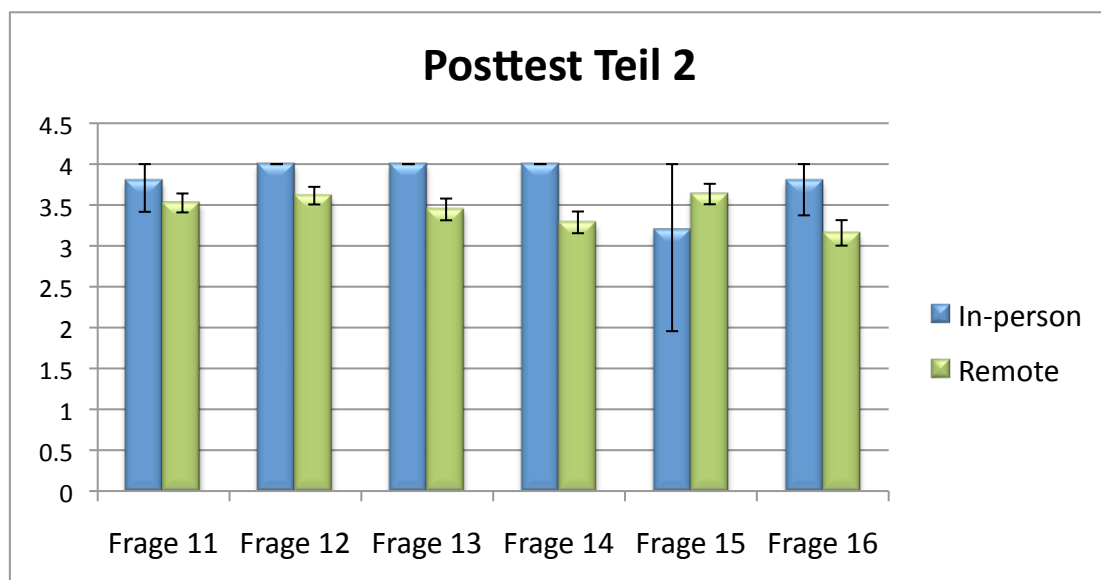


Bei der System Usability Scale verhielt es sich ähnlich wie bei der Single Ease Question. Mit einem Durchschnittswert von 80 zu 78 gibt es zwischen den Methoden fast keinen Unterschied. Dies obwohl beim In-person Test nur fünf Personen teilgenommen haben und die Remote Teilnehmer die Fragen ohne Aufsicht oder ohne eventuelle Beeinflussung eines Moderators beantworteten. Einen signifikanteren Unterschied hingegen gibt es beim zweiten Teil des Posttests. Daher wird auf diesen Test im anschließenden Kapitel näher eingegangen.

6.1.3 Zweiter Posttest-Fragebogen

Die Tabelle 8 stellt die einzelnen Fragen des zweiten Posttest-Fragebogens dar. Jede Frage konnte auf einer Skala von null bis vier bewertet werden, wobei vier die positivste Antwort war.

Tabelle 8: Resultat des zweiten Teils des Posttest-Fragebogens



Die Teilnehmer des In-person Usability-Tests haben den zweiten Teil des Posttest-Fragebogens durchschnittlich besser bewertet. Einzig die Frage 15 schnitt bei dieser Methode schlechter ab. Dieser Unterschied wurde durch einen einzelnen Teilnehmer des In-person Tests generiert, der antwortete, sich sehr unter Druck gefühlt zu haben, die Aufgaben richtig zu lösen. Da nur ein Teilnehmer negativ geantwortet hat, ist das Konfidenzintervall auch entsprechend hoch und es handelt sich nicht um einen signifikanten Unterschied.

Die Fragen des zweiten Posttest-Fragebogens bezogen sich auf die Konzeption des Tests, welcher vom Moderator konzipiert wurde. Es ist daher möglich, dass die Teilnehmer beeinflusst wurden. Sie kannten den Moderator persönlich und beantworteten die Fragen vielleicht positiver, um dem Moderator nicht zu nahe zu treten.

6.1.4 Fazit zur Hypothese 1

Die beiden Indikatoren SEQ und SUS konnten die Hypothese 1 nicht stützen und sprechen eher gegen eine Veränderung der subjektiven Einstellung während eines Remote oder In-person Tests. Trotz eines sehr unterschiedlichen Verhältnisses der Teilnehmerzahl pro Test (5 zu 101) liegen die durchschnittlichen Werte erstaunlich nahe beieinander. Dies spricht dafür, dass auch bei einem Discount Usability-Test mit fünf Teilnehmern statistische Daten mit der Single Ease Question oder der System Usability Scale erhoben werden können, welche einen Richtwert für die Einfachheit der Aufgaben oder für die Usability der Benutzeroberfläche liefern.

Der einzig feststellbare Unterschied bei der subjektiven Einstellung zeigte sich bei dem dritten Indikator, dem zweiten Posttest-Fragebogen. Die positivere Bewertung der In-person Test Teilnehmer lässt sich dadurch erklären, dass Moderator und Teilnehmer sich persönlich kannten. Die Teilnehmer hatten möglicherweise das Gefühl, den Moderator, der den Test konzipiert hatte, zu bewerten und gaben deshalb bessere Noten.

Nach der Evaluation der beiden Nutzertests kann die Hypothese 1 wie folgt angepasst werden:

Hypothese 1 b: *Die subjektive Einstellung der Testpersonen während eines unmoderierten Remote Usability-Tests ist identisch mit der Einstellung während eines In-person Usability-Test, solange bei Fragen keine Verbindung zu den Fähigkeiten des Moderators erstellt werden kann, falls eine persönliche Beziehung zwischen dem Teilnehmer und dem Moderator besteht.*

Abschliessend lässt sich zusammenfassen, dass eine Beeinflussung während des In-person Tests dadurch minimiert werden könnte, dass entweder Moderator und Teilnehmer sich nicht persönlich kennen oder dadurch, dass keine Fragen gestellt werden, die mit den Fähigkeiten des Moderators in Verbindung gebracht werden können.

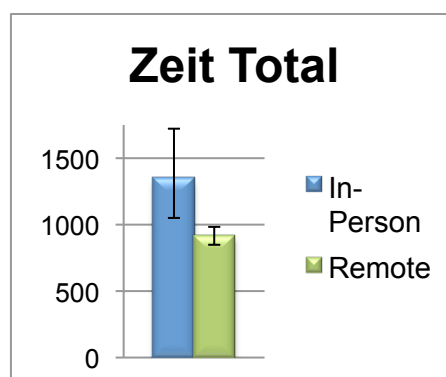
6.2 Überprüfung des Verhaltens

Bei der zweiten Hypothese sollte nicht mehr die subjektive Beantwortung von Fragen untersucht werden, sondern ob und wie die Teilnehmer ihr Verhalten bei den jeweiligen Testmethoden verändern. Folgende Aussage sollte geprüft werden: „Das Verhalten einer Testperson während eines unmoderierten Remote Usability-Tests ist anders und realistischer als während eines klassischen Nutzertests.“

6.2.1 Benötigte Zeit

Ein erster Indikator für diese Hypothese war die benötigte Zeit. Gemäss den Empfehlungen von Sauro³¹ soll für die Darstellung von Aufgabenzeiten eine Transformation verwendet werden. Dies ist nötig, da Aufgabenzeiten nicht symmetrisch um den Durchschnitt verteilt sind, sondern die Kurve links flacher abfällt. Das heisst, es gibt nur einzelne sehr hohe Werte. Die Werte mussten zuerst transformiert werden, damit sie einer normalen Verteilung entsprechen. Die Balken der Tabellen 9 und 10 stellen daher nicht den arithmetischen Durchschnitt, sondern den etwas tieferen geometrischen Durchschnitt³² dar und das Konfidenzintervall ist nicht symmetrisch. Durch die Verwendung des geometrischen Durchschnitts werden die Auswirkungen von unerwartet hohen Zeiten minimiert, weshalb die Daten nicht von solchen Ausreissern bereinigt werden mussten (Albert, Tullis, Tedesco, 2010: 134 f.). Einzig die Zeiten von Teilnehmer Nr. 94 des Online Tests wurden nicht berücksichtigt, da Loop11 diese Daten nicht lieferte. Zuerst wird in der Tabelle 9 die durchschnittliche Gesamtzeit nach Methode verglichen. Auf der y-Achse wird die Zeit in Sekunden angegeben.

Tabelle 9: Durchschnittlich benötigte Zeit für den gesamten Test



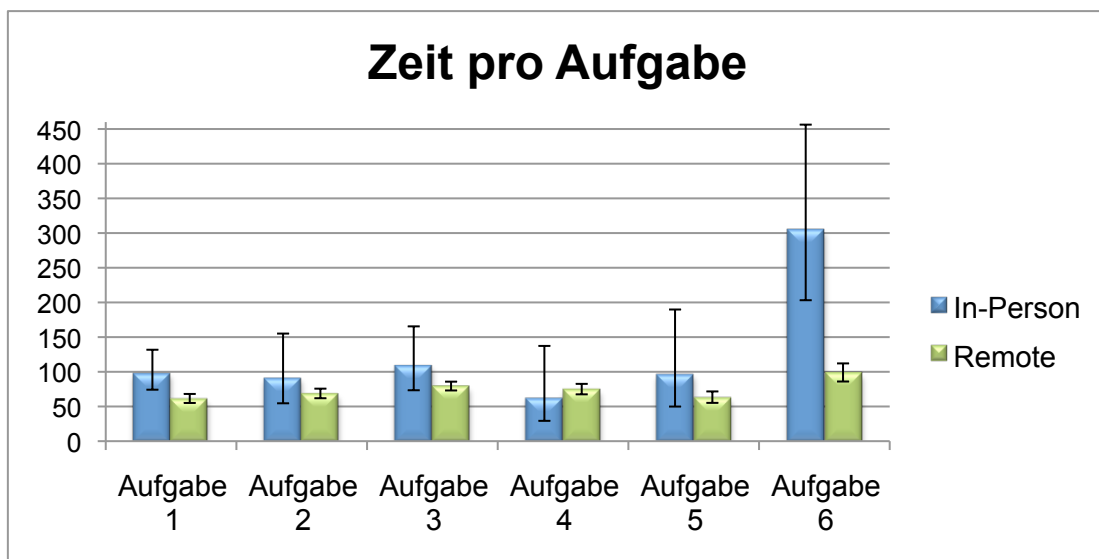
³¹ SAURO, Jeff. Graph and calculator for confidence intervals for task times. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 06.02.2006.
http://www.measuringusability.com/time_intervals (Zuletzt konsultiert: 26.06.2013)

³² Alle n Werte werden multipliziert und anschliessend wird aus dem Produkt die n-te Wurzel gezogen.

Zwischen der Gesamtdauer eines In-person Tests und eines Remote Tests besteht ein signifikanter Unterschied. Ein In-person Test dauerte eindeutig länger. Ein Grund für die längere Dauer eines In-person Tests ist der mündliche Austausch vor und nach den Aufgaben mit dem Moderator. Trotzdem gab es bei den Teilnehmern des Remote Tests einige sehr lange Zeiten. Diese einzelnen, überdurchschnittlich hohen Zeiten wirken sich aber nicht auf den Durchschnittswert aus, da das geometrische Mittel verwendet wurde, welches die Auswirkungen von Ausreißern minimiert. Da diese Teilnehmer nicht überdurchschnittlich viele Seiten besuchten, waren diese Teilnehmer wohl während einer gewissen Zeit mit einer anderen Tätigkeit beschäftigt.

Aber benötigten die Teilnehmer des In-person Tests zum Lösen einer Aufgabe wirklich mehr Zeit oder wird nur die Gesamtzeit durch den mündlichen Austausch beeinflusst? Wenn die Tabelle 10 genauer betrachtet wird, fällt auf, dass ausser bei der Aufgabe 4 die Teilnehmer des In-person Test durchschnittlich immer länger gebraucht haben. Besonders auffällig ist dieser Unterschied bei den Aufgaben 1 und 6, da sich nur bei diesen Aufgaben die Konfidenzintervalle nicht überschneiden.

Tabelle 10: Durchschnittlich benötigte Zeit pro Aufgabe



Dass die Teilnehmer des Remote Tests bei Aufgabe 4 länger brauchten, lässt sich bei genauerer Analyse des Testaufbaus erklären. Bei der Aufgabe 4 handelte es sich um eine Folgeaufgabe der Aufgabe 3. Teilnehmer, die mit Morae aufgezeichnet wurden, konnten direkt mit dem Bildschirm aus Aufgabe 3 weiterarbeiten. Mit Loop11 war es nicht möglich, eine neue Aufgabe beginnend mit einem anderen Link als der Startseite von Swissbib zu definieren. Dies lag daran, dass die von Swissbib generierten Links zu lange waren, um sie als Startpunkt festzulegen. Daher mussten die Teilnehmer des Remote Tests das Buch erneut suchen, was zu einer etwas längeren Dauer geführt hat.

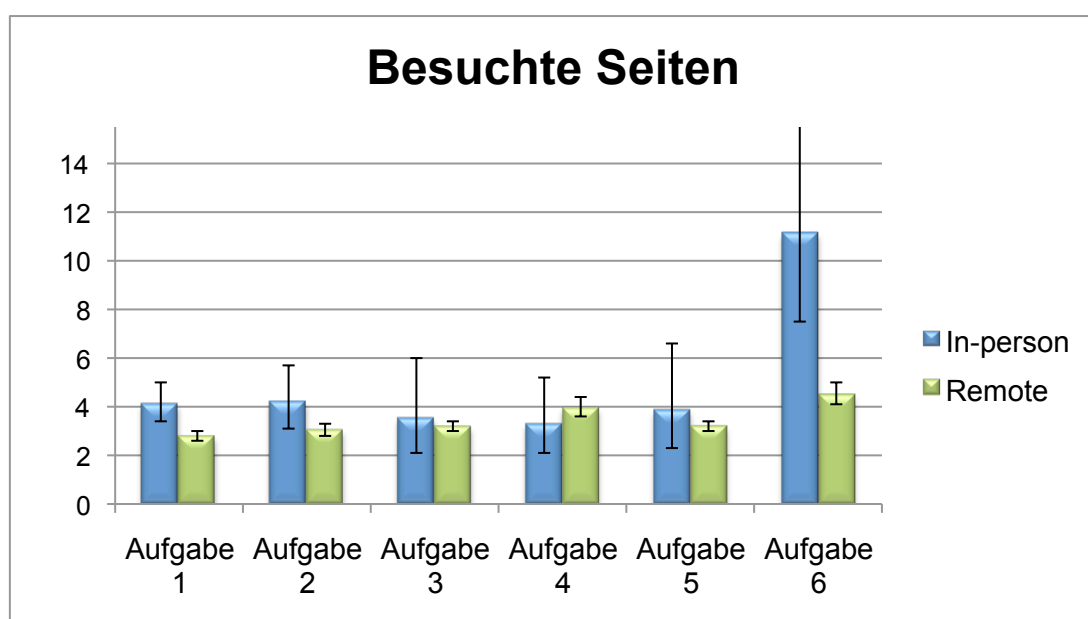
Die Unterschiede bei Aufgabe 2,3 und 5 sind nicht signifikant und lassen sich wahrscheinlich auf das Thinking Aloud zurückführen. Das laute Sprechen und Erklären von Handlungen verlängert die durchschnittliche Dauer hier aber nur minimal.

Bei Aufgabe 1 und besonders bei Aufgabe 6 verlängerten sich die Aufgabenzeiten beim In-person Nutzertest aber signifikant. Dies liegt daran, dass diese beiden Aufgaben einen anderen Charakter haben als die restlichen. Bei Aufgabe 1 mussten sich die Teilnehmer einen ersten Eindruck von Swissbib machen, indem sie ein beliebiges Buch suchten. Auch die letzte Aufgabe hatte einen eher explorativen Charakter. Es ging nicht darum, eine präzise Antwort im Katalog zu suchen. Das Ziel war es, die Funktionen eines Benutzerkontos zu erkunden. Beim Online Test liessen sich die Benutzer weniger Zeit, um sich mit diesen Aufgaben zu beschäftigen. Dies kann daran liegen, dass die In-person Teilnehmer durch die Beobachtung eines Moderators mehr motiviert werden, die Aufgabe möglichst vollständig zu lösen.

6.2.2 Besuchte Seiten

Der zweite Indikator zur Überprüfung des Verhaltens war die Anzahl besuchter Seiten pro Aufgabe. Die Sichtung der Daten zeigte, dass, genau wie bei den Aufgabenzeiten, die Daten nicht normalverteilt waren. Daher wurde ebenfalls eine Transformation vorgenommen und das geometrische Mittel berechnet. Durch diese Vorgehensweise mussten die Daten nicht von Ausreißern bereinigt werden. Einzig der Teilnehmer Nr. 94 des Online Tests konnte erneut nicht berücksichtigt werden, da einige Daten fehlten.

Tabelle 11: Durchschnittliche Anzahl besuchter Seiten



Die Tabelle 11 zeigt die durchschnittliche Anzahl besuchter Seiten für jede Aufgabe. Die Verteilung der besuchten Seiten stimmt erwartungsgemäss mit der Verteilung der Aufgabenzeiten überein. Wieder zeigt einzig die Aufgabe 4 einen grösseren Wert für den Remote Test an. Dies lässt sich ebenfalls dadurch erklären, dass es sich hier um eine Folgeaufgabe handelte und die Teilnehmer mit Loop11 das Dokument erneut suchen mussten, bevor sie mit der Aufgabe beginnen konnten.

Die Teilnehmer des In-person Tests haben bei Aufgabe 2, 3 und 5 nicht nur mehr Zeit benötigt, sondern auch etwas mehr Seiten besucht. Die Unterschiede bei diesen drei Aufgaben sind jedoch nicht signifikant. Erneut liefern nur Aufgabe 1 und Aufgabe 6 aussagekräftige Unterschiede. Wie bereits erwähnt, handelt es sich hier um Aufgaben ohne klares Ziel, bei denen die Teilnehmer frei sind sich auf der Seite zu bewegen, um Informationen zu sammeln. Besonders bei der Aufgabe 6 suchten die Teilnehmer des Online Tests auf wenigen Seiten nach Informationen zum Benutzerkonto. Sie begnügten sich oftmals damit, sich anzumelden und zusätzlich noch eine Hilfe-Seite oder das Konto zu konsultieren. Es ist möglich, dass die Teilnehmer des In-person Tests durch die Anwesenheit eines Moderators motiviert wurden, ausführlicher zu suchen.

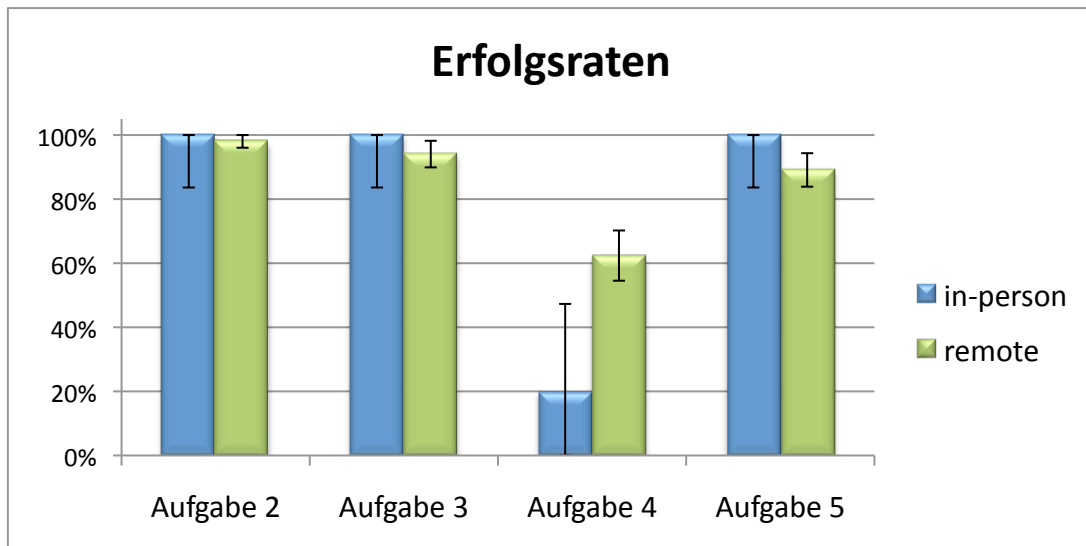
6.2.3 Erfolgsrate

Der dritte und letzte Indikator für die Hypothese 2 über das Verhalten der Teilnehmer ist die Erfolgsrate. Die Aufgaben 1 und 6 hatten keine Erfolgsrate, da es keine klar richtige oder falsche Antwort auf diese Fragen gab. Die Frage 1 beschäftigte sich mit dem ersten Eindruck und für die Frage 6 war es möglich, viele mögliche Funktionen des Benutzerkontos zu beschreiben. Es war daher nicht möglich zu bestimmen, ab wie vielen gefundenen Funktionen die Aufgabe richtig gelöst wurde.

Anhand der Multiple-Choice-Frage am Ende der Aufgaben wurde entschieden, ob ein Teilnehmer die Aufgabe richtig oder falsch gelöst hatte. Die Teilnehmer hatten bei jeder Aufgabe die Möglichkeit anzukreuzen, dass sie die Aufgabe nicht lösen konnten. Sowohl beim In-person als auch beim Remote Test wurde diese Option jedoch sehr selten verwendet. Beim In-person Test wurde umgehend bemerkt, wenn ein Teilnehmer riet und aus Zufall die richtige Antwort ankreuzte, obwohl er die Aufgabe nicht lösen konnte. Beim Remote Test war es jedoch schwierig zu beurteilen, ob ein Teilnehmer die Antwort tatsächlich gefunden oder nur geraten hatte. Besonders bei Teilnehmern, die sehr kurz oder sehr lange brauchten, um eine Aufgabe zu lösen, wurden einzelne Stichproben gemacht und Navigationswege überprüft. In wenigen Fällen mussten Korrekturen gemacht werden. Diese Tatsache zeigt, dass es schwierig ist, die exakten

Erfolgsraten von unmoderierten Remote Tests durch Multiple-Choice-Fragen zu bestimmen. Alternativ könnte eine offene Frage am Ende einer Aufgabe gestellt werden, jedoch ist dies auch problematisch, da unkonzentrierte Benutzer die Antworten wieder vergessen könnten.

Tabelle 12: Durchschnittliche Erfolgsraten pro Aufgabe

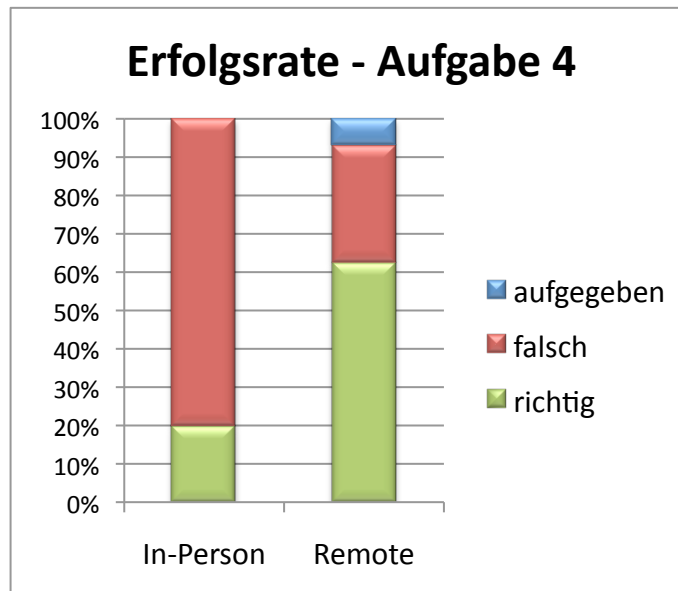


Die Tabelle 12 stellt dar, wie viel Prozent der Teilnehmer der jeweiligen Testmethode die Aufgaben richtig lösen konnten. Ausser bei der Aufgabe 4 haben beim im In-person Test alle fünf Teilnehmer die Lösungen gefunden. Beim In-person Test sinkt die Erfolgsrate bei jeder Aufgabe kontinuierlich. Dies entspricht dem Anstieg der Schwierigkeit mit jeder weiteren Aufgabe. Die Unterschiede zwischen In-person und Remote Test sind zwar nicht signifikant, jedoch ist die Erfolgsrate ausser bei Aufgabe 4 immer etwas tiefer. Diese Tatsache hängt wohl damit zusammen, dass sich Teilnehmer eines In-person Tests, wie in den vorangehenden Kapiteln gesehen, durchschnittlich mehr Zeit nehmen und mehr Seiten besuchen als während eines Online Tests. Da sich Teilnehmer des unmoderierten Remote Tests alleine vor dem Bildschirm befinden, bemühen sie sich weniger lange nach der richtigen Antwort zu suchen und werden nicht von einem Moderator motiviert.

Auffällig ist, dass die Aufgabe 4 nicht ins Bild passt. Das Ziel dieser Aufgabe war es, zu überprüfen, ob ein bestimmtes Hörbuch ausleihbar sei. Nur einer der fünf getesteten Personen beim In-person Test erkannte richtig, dass das Dokument bereits von einer anderen Person ausgeliehen war. Alle anderen dachten, dass das Hörbuch ausleihbar sei. Auch wenn sie sich nicht sicher waren, gab keine der Personen an, die Aufgabe nicht lösen zu können. Beim Remote Test haben rund 62% die Aufgabe richtig gelöst,

31% dachten das Dokument sei ausleihbar und 7% gaben an, die Aufgabe nicht lösen zu können (siehe Tabelle 13).

Tabelle 13: Erfolgsrate der Aufgabe 4



Diese im Vergleich zu den anderen Aufgaben schlechten Erfolgsraten können durch unterschiedliche Faktoren begründet sein. Einerseits nutzten viele Teilnehmer den Internet Explorer 8, beim welchem die Verfügbarkeit nicht angezeigt wird (siehe Kapitel 5.8, Problem Nr. 18). Dadurch waren sie gezwungen auf den Katalog der besitzenden Bibliothek zurückzugreifen. Wenn ein Teilnehmer jedoch die Verfügbarkeit auf dem Katalog der besitzenden Bibliothek überprüfte, fand er widersprüchliche Informationen. Obwohl ein Fälligkeitsdatum angegeben war, stand das Dokument sei „ausleihbar“ (Abb. 38). Die Teilnehmer folgerten dadurch fälschlicherweise, das Buch sei nicht ausgeliehen. Der einzige Teilnehmer des In-person Tests, der die Frage richtig beantworten konnte, war der Bibliothekar.

Abbildung 38: Anzeige der Verfügbarkeit - Lokaler Katalog

fällig am	Anzahl Vorm.	Exemplarstatus
02/08/13		ausleihbar

Wenn sie aber die Information auf Swissbib gefunden hätten, hätten sie gesehen, dass das Hörbuch bereits ausgeliehen war, denn eigentlich zeigt Swissbib die Verfügbarkeit eines Dokuments gut mit einem Symbol an (Abb. 39).

Abbildung 39: Anzeige der Verfügbarkeit - Swissbib



Andererseits wird, auch wenn nicht der Internet Explorer 8 genutzt wird, die Verfügbarkeit bei Swissbib nicht direkt, sondern erst beim Anklicken von „Informationen zum Bestand“ angezeigt. Häufig wurde diese Option gar nicht aufgerufen und die Teilnehmer klickten direkt auf den Link „zum Bestand“, welcher auf den lokalen Bibliothekskatalog verweist (siehe Kapitel 5.7, Problem Nr. 15).

Dass die Teilnehmer des Remote Tests bei der Aufgabe 4 durchwegs besser abschnitten, liegt daran, dass die Teilnehmer aus dem Bereich Information und Dokumentation kamen. 75% aller Personen hatten sehr viel Erfahrung oder waren sogar Experten auf dem Gebiet. Deshalb konnten sie die widersprüchlichen Informationen auf dem lokalen Katalog besser interpretieren.

6.2.4 Fazit zur Hypothese 2

Aus den untersuchten Indikatoren „benötigte Zeit“ und „besuchte Seiten“ liessen sich tatsächlich Unterschiede im Verhalten feststellen. Einige Ausreisser zeigen, dass Teilnehmer während eines Online Tests von ihrer Umgebung abgelenkt werden und dadurch länger brauchen können um eine Aufgabe zu lösen. Der Einfluss dieser Zeiten kann durch die Verwendung des geometrischen Mittels minimiert werden und so können trotzdem realistische Durchschnittswerte mit dem Remote Test berechnet werden.

Bei den Aufgaben mit einem klaren Ziel (Aufgaben 2 bis 5) waren die Unterschiede nicht signifikant. Nur bei explorativen Aufgaben ohne konkretes Ziel benötigten die Teilnehmer des In-person Tests durchschnittlich mehr Zeit zum Lösen einer Aufgabe und besuchen auch mehr Seiten. Für diese Aufgaben wurden keine Erfolgsraten bestimmt, so dass keine Aussage über diesen Indikator getroffen werden kann. Die längere Dauer lässt sich teilweise durch das Thinking Aloud und durch den mündlichen Austausch mit dem Moderator erklären. Da aber die Teilnehmer auch mehr Seiten besuchten, ist dies nicht der einzige Grund. Bei Aufgaben ohne klares Ziel betrieben beim In-person Test die Benutzer mehr Aufwand die Aufgaben zu lösen. Aufgaben, die einen explorativen Charakter haben, eignen sich daher nur bedingt für einen unmoderierten Remote Test, da Benutzer sich zu wenig damit auseinandersetzen.

Diese Tatsachen stützen die Hypothese 2. Die Teilnehmer eines Remote Tests können einerseits von der natürlichen Umgebung abgelenkt werden, was bei einigen wenigen Teilnehmern zu sehr hohen Aufgabenzeiten führt. Durchschnittlich wenden Personen während eines Remote Tests aber weniger Zeit und Aufwand auf, um sich auf einer Webseite zurechtzufinden und um eine Antwort zu finden. Dies trifft aber nur zu, falls es sich um Aufgaben ohne konkretes Ziel handelt. Dies kann daran liegen, dass bei einem In-person Test die Teilnehmer durch einen Moderator beeinflusst und motiviert werden die Aufgaben möglichst richtig zu lösen. Wenn Teilnehmer eines Remote Tests brauchen scheinbar ein konkretes Ziel, um klare Richtlinien zu haben und um die Aufgaben gut zu lösen.

Ob dieses Verhalten auch tatsächlich realistischer ist, lässt sich durch diese Arbeit nicht überprüfen. Bei einer echten Nutzungssituation hat ein Benutzer ein persönliches Interesse, eine bestimmte Information zu finden und wird vielleicht deshalb mehr Zeit investieren. Die Hypothese 2 lässt sich folgendermassen anpassen:

Hypothese 2 b: *Testpersonen eines unmoderierten Remote Usability-Tests verwenden durchschnittlich weniger Zeit und besuchen weniger Seiten bei Aufgaben ohne klarem Ziel als Teilnehmer eines Discount Nutzer-tests.*

Die Resultate bei Aufgaben mit einem klaren Ziel und einer klaren Antwort waren zu wenig signifikant, als dass tatsächlich von einer unterschiedlichen Beeinflussung durch die Testmethode ausgegangen werden kann.

6.3 Überprüfung der Effektivität

Das primäre Ziel eines Usability-Tests ist, wie bereits erwähnt, die Identifikation von Usability-Problemen. Die dritte Hypothese „Ein unmoderierter Remote Usability-Test ist genauso effektiv wie ein In-person Discount Usability-Test.“ sollte analysiert werden, indem in jedem Test die gefundenen Usability-Probleme gezählt und nach Schweregrad geordnet werden. Eine ausführliche Beschreibung aller Probleme findet sich unter Kapitel 5. Informationen darüber, mit welcher Methode die Schwierigkeiten gefunden wurden, sind im Anhang H zu finden.

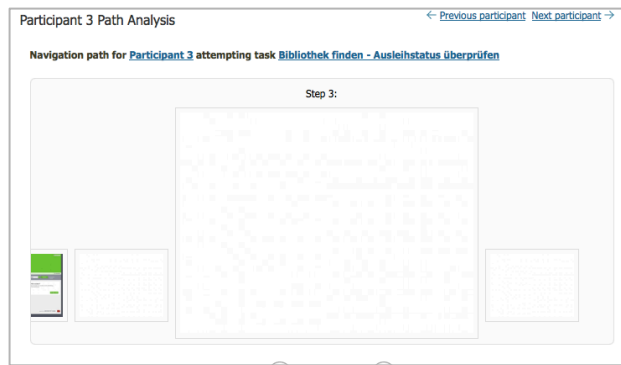
Für den In-person Test wurden zur Identifikation der Probleme die mit Morae aufgezeichneten Daten gesichtet. Dazu wurden die Tonaufnahmen nicht, wie von Schweibenz und Thissen (2003: 162) empfohlen, transkribiert, da dies zeitaufwendig ist und nicht dem Sinn eines Discount Usability-Tests entspricht. Die Bildschirm- und Videoaufnahmen wurden visioniert und es wurden Probleme sowie aussagekräftige Kommentare von Teilnehmern notiert. Mit Hilfe dieser Daten wurden die Usability-Probleme identifiziert und entsprechend ihrem Schweregrad geordnet. Je häufiger ein Problem auftrat und je grösser die Irritation beim Benutzer war, desto schwerwiegender wurde es gewichtet.

Albert, Tullis und Tedesco (2011: 169 f.) empfehlen drei Techniken, die dabei helfen sollen, Usability-Probleme durch einen Remote Test zu identifizieren:

- Sich auf Aufgaben konzentrieren, mit welchen Teilnehmer die meisten Probleme hatten
- Fehler der Teilnehmer analysieren
- Kommentare der Teilnehmer analysieren

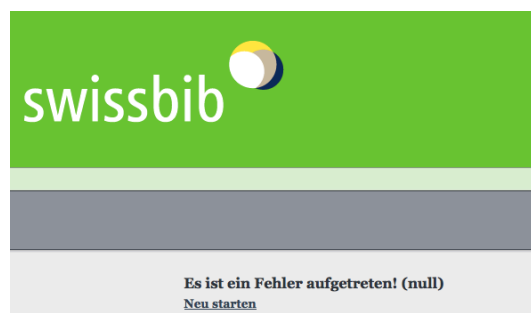
Besonders die Kommentare der Teilnehmer zu Beginn oder am Ende des Tests waren sehr hilfreich. Aber die Software Loop11 zeichnet noch weitere Daten auf, welche genutzt werden können, um Usability-Probleme zu identifizieren. Die Funktion „Participant path analysis“ soll darstellen, welche Seiten ein bestimmter Benutzer besucht hat. Gemäss Albert, Tullis und Tedesco (2011: 169) soll besonders diese Funktion eines Remote Tools dabei helfen, konkrete Probleme zu identifizieren. Die Benutzer, die eine Aufgabe nicht, nur nach langer Zeit oder durch den Besuch von überdurchschnittlich vielen Seiten lösen konnten, geben am meisten Auskunft darüber, wo mögliche Schwierigkeiten sein könnten. Genau diese Funktion arbeitete beim Benutzertest von Swissbib aber leider nicht sachgemäss (Abb. 40).

Abbildung 40: Problem mit der Funktion „Path-Analysis“



Loop11 hatte Probleme bei der Analyse der Links von Swissbib. Die besuchten Seiten wurden zwar aufgelistet, konnten aber nicht in einer Vorschau angezeigt werden. Teilweise konnten die Seiten durch das Anklicken des Links trotzdem aufgerufen werden. In vielen Fällen war dies aber nicht möglich und es erschien eine Fehlermeldung (Abb. 41). Loop11 konnte das Problem bis zum Abschluss der Arbeit nicht lösen und der Grund für dieses Problem konnte nicht genau identifiziert werden. Es lag wohl teilweise auch an der Art, wie Swissbib Hyperlinks generiert, so dass diese nach einer bestimmten Zeit ablaufen.

Abbildung 41: Fehlermeldung



Auch Klickstream-Daten, die die Wege aller Teilnehmer zusammenfassen, sind hilfreich beim Erkennen von Schwachstellen. Sie können aufzeigen, von welchen Seiten am häufigsten eine Aufgabe abgebrochen wird oder welche Wege am häufigsten gegangen werden. Beim Test auf Swissbib mussten die Benutzer aber meistens Recherchen in einem Suchfeld durchführen. Solche Eingaben sind sehr individuell und daher wird jedesmal ein anderer Link generiert. Auch wenn zwei Benutzer dieselbe Suche ausführten, wurden die Seiten von Swissbib nicht zusammengeführt. Die Analyse der Klickstream-Daten war daher kaum möglich (siehe Abb. 42).

The screenshot shows a web browser window with the address bar displaying "http://www.sensib.it". The main content area is filled with a dense network of overlapping, semi-transparent rectangular panels. Each panel represents a web page and contains various elements such as navigation menus, search bars, and text. The panels are interconnected by a complex web of lines, representing the links between the pages. The overall effect is a visual representation of a large, interconnected network of web pages.

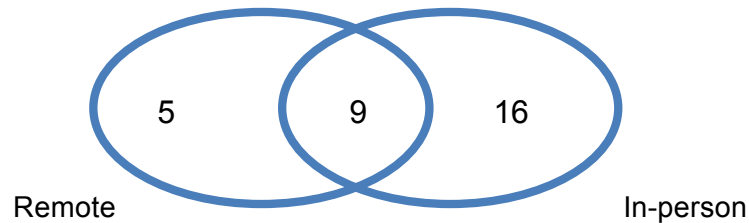
Abbildung 43: Heatmap einer spezifischen Seite aus Loop11

58

6.3.1 Anzahl und Schweregrad der identifizierten Usability-Probleme

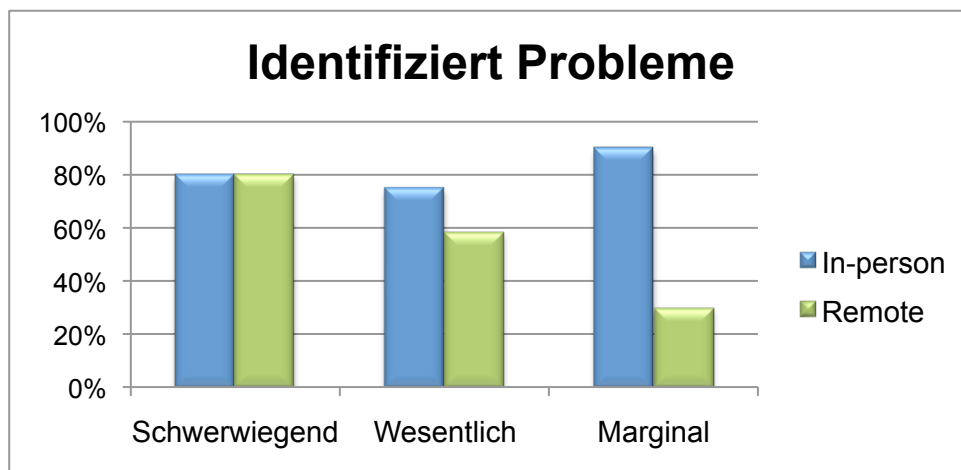
Insgesamt konnten durch die Analyse der Tests 27 Usability-Probleme identifiziert werden, davon waren 5 schwerwiegend, 12 wesentlich und 10 marginal. Mit dem In-person Test konnten 22 von 27 und mit dem Remote Test 14 von 27 Problemen identifiziert werden (Abb. 45).

Abbildung 45: Anzahl identifizierte Usability-Probleme



Hier scheint es, als ob der Discount Nutzertest trotz wenigen Teilnehmern effektiver ist als der Online Test. Jedoch ist nicht nur die Anzahl der Probleme sondern vor allem der Schweregrad der gefundenen Probleme massgebend. Die Tabelle 14 zeigt, welche Methode wie gut die Probleme jedes Schweregrads identifizieren konnte.

Tabelle 14: Anteil identifizierter Probleme je Schweregrad



Diese Grafik zeigt, dass der In-person und der Remote Test bei den wichtigsten, den schwerwiegenden Problemen identisch abschnitten. Beide konnten 80% der schwerwiegenden Nutzerprobleme finden. Bei den wesentlichen Unterschieden war der Remote Test etwas weniger effektiv. Der Grund für das insgesamt schlechtere Abschneiden des Remote Tests liegt vor allem bei den marginalen Problemen. Gerade mal drei von zehn aus dieser Kategorie wurden gefunden. Was sind die möglichen Ursachen für diese Unterschiede?

6.3.2 Beim Remote Test nicht identifizierte Probleme

Ein Grund für das rein zahlenmässig schlechtere Abschneiden des Remote Tests ist einerseits, dass die von Loop11 gelieferten Daten nicht erwartungsgemäss analysiert werden konnten. Wie bereits im vorangehenden Kapitel beschrieben wurde, arbeitete die Funktion „Participant path analysis“ nicht vollständig und der genaue Weg eines Teilnehmers konnte nicht in jedem Fall nachkonstruiert werden. Durch die sehr unterschiedlichen Suchanfragen und durch das mangelnde Zusammenführen von gleichen Suchanfragen bei Loop11 lieferten weder die Heatmaps noch die Klickstream-Daten wirklich gute Hinweise auf Probleme. Ohne diese Schwierigkeiten hätte der unmoderierte Remote Test vielleicht einige Probleme mehr aufgedeckt. So hätte beispielsweise die fehlende Anzeige der Verfügbarkeit mit dem Internet Explorer 8 durch die Analyse der Wege der Teilnehmer entdeckt werden können. Auch die Fehlversuche, ein Dokument mit dem Bild oder dem Merksymbol zu öffnen, hätten durch eine entsprechende Heatmap gezeigt werden können.

Die Gründe für das schlechtere Resultat des Remote Tests sind aber nicht nur technischer Natur. Für den unmoderierten Online Nutzertest wurde ein Tool verwendet, das keine Video- oder Tonaufnahmen der Teilnehmer speichert. Dadurch können Gefühle und Stimmungen der Teilnehmer nicht registriert werden. So wurde nicht bemerkt, wann ein Benutzer irritiert war, ausser er nutzte die Kommentarfelder am Anfang oder Ende des Tests, um seine Irritation zum Ausdruck zu bringen. Dies war meistens nur der Fall, wenn es sich um schwerwiegende Probleme handelte. Der Mangel an Feedback wird offensichtlich, wenn es sich um Probleme mit der Bezeichnung von Menüoptionen oder Buttons geht. So wurde im Remote Test beispielsweise nicht entdeckt, dass die Bezeichnungen „Zum Bestand“ oder „Temporäre Merkliste“ ein Problem für die Benutzer darstellen (Kapitel 5, Probleme Nr. 16 und 27). Jedoch erlaubt der Remote Test, verwirrende Navigationsstrukturen aufzuzeigen, wie es das Beispiel mit dem Unterschied zwischen „Zum Bestand“ oder „Informationen zum Bestand“ zeigt (Problem Nr. 15). Weiter hat sich in den schriftlichen Kommentaren keiner der Teilnehmer des Remote Tests konkret über die Grösse der Menünavigation, den unklaren Link zur erweiterten Suche oder die Anzeige des Marc-Formats beschwert, wodurch diese eher marginalen Probleme nicht gefunden werden konnten. Durch weitere Kommentarmöglichkeiten am Ende jeder Aufgabe hätten diese Probleme vielleicht besser identifiziert werden können. Die Bearbeitung von vielen freien Kommentarfeldern ist aber mit zusätzlicher Arbeit verbunden und wie Bolt und Tulathimutte (2010: 123) festhalten, sind spontane Thinking Aloud Kommentare nicht zu vergleichen mit überlegten schriftlichen Kommentaren.

Neben den fehlenden spontanen Rückmeldungen der Benutzer ist auch die mangelnde Aufzeichnung des Bildschirms mit den Mausbewegungen ein Hindernis. Der Remote Test konnte vor allem Probleme mit dynamischen Elementen, die keinen neuen Link generieren, nicht aufzeigen. Dadurch wurden die Schwierigkeiten mit der Navigation innerhalb der Hilfe (die keine neuen Links generiert) oder mit der Autovervollständigung nicht gefunden (Probleme Nr. 2,3, und 21). Ein Benutzer machte auf Loop11 zwar den Kommentar „Warum werden die Vorschläge nicht übernommen?“, wobei er wohl auf ein Problem mit der Autovervollständigung hinwies, jedoch konnte wegen der fehlenden Bildschirmaufzeichnung nicht nachvollzogen werden, was genau er mit dieser Aussage gemeint hatte. Dieses Beispiel zeigt auch, dass schriftliche Kommentare zwar sehr hilfreich sein können, jedoch nicht immer verständlich sind, wenn der Benutzer nicht bei seinen Handlungen beobachtet werden kann. Inzwischen gibt es Tools, die Bildschirm- oder Tonaufnahmen liefern. Diese könnten bei Unklarheiten konsultiert werden und helfen, die Benutzer besser zu verstehen. Jedoch könnte so die Vorbereitung technisch anspruchsvoller und die Auswertung zeitintensiver werden.

6.3.3 Beim In-person Test nicht identifizierte Probleme

Obwohl der Discount Test deutlich mehr Nutzungsprobleme aufzeigen konnte, konnte er fünf nicht identifizieren, die mit dem Remote Test gefunden wurden. Zwei dieser unentdeckten Probleme konnten damit erklärt werden, dass keiner dieser Teilnehmer einen Schreibfehler beim Abtippen machte. Dadurch wurden Probleme bei Null-Treffermengen nicht identifiziert (Kapitel 5, Probleme Nr. 4 und 6). Grund hierfür kann sein, dass beim In-person Test zu wenige Personen teilnahmen, so dass das relativ seltene Problem überhaupt auftauchte oder dass sich die Teilnehmer während des In-person Tests beim Abtippen mehr konzentrierten und deshalb keine Schreibfehler machten.

Der Mangel an Daten könnte auch mit der Grund sein, weshalb der unklare Aufbau der Hilfethemen durch den In-person Test nicht erkannt wurde (Problem Nr. 20). Zudem nahmen sich die fünf Teilnehmer mehr Zeit, um sich in den Hilfethemen zurechtzufinden, wodurch sie die gesuchten Informationen fanden. Durch den In-person Test wurde auch nicht entdeckt, dass das Symbol für die Informationen über eine Bibliothek zu klein ist (Problem Nr. 17). Der Grund dafür ist, dass diese Funktion im Verlaufe des Tests eigentlich gar nicht benutzt werden musste. Jedoch haben zwei sehr erfahrene Benutzer von Swissbib beim Online Test auf diese Schwierigkeit hingewiesen. Nicht identifiziert wurde ebenso das eventuelle Problem mit der Verarbeitung von Umlauten. Dies kann daran liegen, dass es ein sehr seltenes Problem ist oder daran, dass es durch die Testumgebung mit Loop11 generiert wurde und kein echtes Nutzerproblem darstellt.

6.3.4 Fazit zur Hypothese 3

Bei den durchgeführten Tests war die Online Testmethode weniger effektiv als die In-person Methode, jedoch nur im Bereich von wesentlichen oder marginalen Schwierigkeiten. Bei den schwerwiegenden Problemen schlossen beide Methoden gleich gut ab. Die Gründe für das schlechtere Resultat sind einerseits bei technischen Schwierigkeiten zu suchen. Aber auch die fehlenden Bild-, Ton- und Bildschirmaufnahmen haben sich auf die Effektivität des Remote Tests ausgewirkt. Der Mangel an diesen Daten hat sich vor allem bei der Identifikation von Problemen mit dynamischen Elementen, die keine Links generieren, gezeigt und bei einzelnen unklaren Bezeichnungen oder anderen verwirrenden Elementen, die die Benutzer aber nicht wirklich daran hindern, die Aufgaben zu lösen. Solche Probleme könnten vielleicht besser aufgedeckt werden, wenn die Benutzer des Remote Tests die Möglichkeit hätten, Kommentare zu jeder Aufgabe zu verfassen. Jedoch sind viele Kommentarfelder auch mit zusätzlichem Aufwand für die Analyse verbunden. Zudem ist anzunehmen, dass schriftliche Kommentare nie dieselben ausführlichen und vielseitigen Informationen bieten, wie spontane Thinking Aloud Kommentare. Dies wurde vor allem bei weniger schlimmen Problemen ersichtlich. Die Hypothese 3 kann wie folgt angepasst werden:

Hypothese 3b: *Ein unmoderierter Remote Usability-Test ist bei wesentlichen und marginalen Nutzerproblemen weniger effektiv als ein In-person Discount Usability-Test.*

Obwohl der In-person Test deutlich besser abschnitt, konnte auch er nicht alle 27 Nutzungsprobleme finden. Dies lag einerseits an der grösseren Konzentration und der vorsichtigeren Arbeitsweise der Teilnehmer. Ein anderer Grund war die kleine Teilnehmerzahl, wie diese bei einem Discount Test üblich ist. Einige Probleme hätten durch eine grössere Anzahl Teilnehmer identifiziert werden können, da es sich teilweise um Probleme handelte, die auch beim Online Test sehr selten vorkamen.

7 Fazit

Für diese Bachelorarbeit wurden zwei fast identische Nutzertests mit dem Metakatalog Swissbib durchgeführt. Einmal in Form eines Discount Tests und einmal über das Internet ohne Moderator. Durch die beiden Tests sollten zum einen die beiden Evaluationsmethoden verglichen werden und zum anderen sollte die Benutzerfreundlichkeit von Swissbib evaluiert werden.

Die Teilnehmer der Tests nahmen Swissbib als sehr benutzerfreundlich wahr, jedoch wurden nur die elementarsten Funktionen des Katalogs getestet. Usability-Probleme konnten vorwiegend im Bereich der Grösse und Lesbarkeit von Texten und Symbolen ausgemacht werden. Anpassungen sind vor allem bei der Menüleiste, der Fassettennavigation oder der Hilfe angebracht. Die Datenverarbeitung könnte ebenfalls noch überarbeitet werden, um die Dublettenbereinigung, die Autovervollständigung und die Sortierung der Resultate zu verbessern. Einige Schwierigkeiten hatten die Benutzer auch bei der Architektur der Seite oder bei einzelnen Bezeichnungen von Schaltflächen. Eine Anpassung an die Sprache der Benutzer wäre sinnvoll und Swissbib sollte für den Internet Explorer optimiert werden. Viel Verwirrung könnte verhindert werden, wenn es den Benutzern möglich wäre, die Dokumente direkt auf Swissbib zu bestellen, zu reservieren und zu verwalten. Um diesem Bedürfnis nachzukommen, ist die Bestellung direkt auf Swissbib bereits in Planung.

Die beiden neueren Usability-Evaluationsmethoden Discount Usability-Test und unmoderierter Remote Usability-Test wurden anhand von mehreren Indikatoren verglichen. Die Ergebnisse zeigten, dass beide Methoden einige Gemeinsamkeiten aufweisen. So wurden die Schwierigkeit der Aufgaben und die Benutzerfreundlichkeit von Swissbib bei beiden Tests gleich eingeschätzt. Bei spezifischen Fragen mit einem genauen Ziel benötigten die Teilnehmer ungefähr gleich lange und besuchten zirka gleich viele Seiten. Für die Identifizierung der schwerwiegendsten Probleme eignen sich die beiden Methoden ebenfalls gleich gut.

Bei Fragen, die mit den Fähigkeiten des Moderators in Verbindung gebracht werden konnten, wurden Unterschiede festgestellt. Teilnehmer des Discount Tests bewerteten diese Fragen besser. Auffallend länger benötigten die In-person Teilnehmer beim Lösen von Aufgaben ohne klares Ziel und sie besuchten bei diesen auch mehr Seiten. Damit die Teilnehmer eines Discount Tests nicht beeinflusst werden, sollten Fragen, die in Verbindung mit dem Moderator gebracht werden können, gemieden werden. Es ist von Vorteil, wenn sich die Teilnehmer und der Moderator nicht persönlich kennen.

Fragen mit explorativem Charakter liefern für Remote Tests nur wenige Daten und sind eher für In-person Tests geeignet. Es ist möglich, dass die Teilnehmer zu Hause oder bei der Arbeit eine weniger grosse Aufmerksamkeitsspanne haben, wenn sie kein konkretes Ziel haben, da sie nicht beobachtet werden.

Mit dem unmoderierten Remote Nutzertest konnten eindeutig weniger mittelschwere und leichte Usability-Problemen identifiziert werden. Technische Schwierigkeiten bei der Analyse können diese schlechtere Effektivität erklären. Zudem eignete sich Swissbib schlecht für die automatische Verarbeitung, da durch die Sucheingaben zu unterschiedliche Links generiert wurden. Um die Effektivität eines unmoderierten Remote Tests zu verbessern, sollten eher Webseiten evaluiert werden, bei welchen sich Benutzer mittels Links durch die einzelnen Seiten navigieren. Dadurch ist die Anzahl Links limitiert und das Auswerten der Daten wird vereinfacht. Mangelnde Kontextinformationen erschwerten es zusätzlich, die Handlungsursachen der Benutzer zu rekonstruieren. Um mehr Kontextinformationen zu erhalten, könnten nach jeder Aufgabe Kommentarfelder eingefügt werden. Dies ist jedoch mit mehr Aufwand bei Auswerten verbunden und schriftliche Kommentare sind nicht so aussagekräftig wie spontane Äusserungen. Neuere Tools erlauben inzwischen Aufnahmen von Bildschirm oder Sprache zu speichern. Diese würden helfen das Handeln der Benutzer besser zu verstehen. Dadurch könnten sich aber möglicherweise technische Schwierigkeiten häufen.

Der unmoderierte Remote Test fand aber einige Probleme, die durch den Discount Test nicht ausgemacht werden konnten. Zwei Gründe könnten erklären, warum der In-person Test nicht alle Probleme gefunden hat. Einerseits ist wohl die kleine Teilnehmerzahl nicht geeignet, alle Nutzerprobleme abzudecken. Andererseits werden Teilnehmer eines In-person Tests nicht vom Alltag abgelenkt und können sich so vielleicht besser konzentrieren. Typische Flüchtigkeitsfehler tauchen daher seltener auf.

Abschliessend lässt sich festhalten, dass sich Discount Usability-Tests und unmoderierte Remote Nutzertests gut eignen, um Einstellungen und das Verhalten von Benutzern zu evaluieren. Jedoch sollten einige Richtlinien berücksichtigt werden, damit Teilnehmer eines Discount Tests nicht durch den Moderator beeinflusst werden und damit die Aufmerksamkeit der Teilnehmer eines Remote Tests bestehen bleibt. Durch die kleine Anzahl Teilnehmer kann ein Discount Test zwar nicht alle, aber sehr viele Probleme identifizieren. Damit ein unmoderierter Online Test möglichst viele Probleme identifizieren kann, ist die Planung sehr wichtig. Nur so können technische Probleme minimiert und genügend Kontextinformationen gesammelt werden. Jedoch hat das Beispiel von Swissbib gezeigt, dass sich nicht jede Seite gleich gut für Online Tests eignet.

Literaturverzeichnis

Swissbib

- BÖTTCHER, Uwe. Swissbib: die ganze Schweiz auf einen Blick. *B.I.T* [online]. 2010, Bd. 13, Nr. 2. S. 157-160. <http://www.b-i-t-online.de/heft/2010-02/nachrichtenbeitraege.pdf> (letzter Zugriff: 08.06.2013)
- SWISSBIB. E-lib.ch/de. In: *Swissbib project wiki* [online]. <http://www.swissbib.org/wiki/index.php?title=E-lib.ch/de> (letzter Zugriff: 09.06.2013)
- SWISSBIB. Project state/de. In: *Swissbib project wiki* [online]. http://www.swissbib.org/wiki/index.php?title=Project_state/de (letzter Zugriff: 09.06.2013)
- SWISSBIB. Über swissbib. In: *Swissbib* [online]. 2013. <http://www.swissbib.ch/TouchPoint/aboutswissbib.do;jsessionid=84599C9761534982B0595C12446FE6D8.worker5?methodToCall=showAboutSwissbib&submenue=nav1&lang=de> (letzter Zugriff: 08.06.2013)
- UNIVERSITÄT BASEL. *Swissbib* [online]. 2013. <http://www.swissbib.ch> (letzter Zugriff: 13.07.2013)

Usability

- BARNUM, Carol M. *Usability testing essentials: ready, set...test!* Amsterdam: Elsevier, 2011. 382 S.
- DUMAS, Joseph S., REDISH, Janice C. *A practical guide to usability testing*. Rev. ed. Exeter: Intellect, 1999. 404 S.
- NIELSEN, Jakob. First rule of usability?: Don't listen to users. In: *Nielsen Norman Group* [online]. 05.08.2001. <http://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/> (letzter Zugriff: 20.06.2013)
- NIELSEN, Jakob. *Usability engineering*. Amsterdam: Morgan Kaufmann, 1993. 362 S.
- SARODNICK, Florian, BRAU, Henning. *Methoden der Usability Evaluation: wissenschaftliche Grundlagen und praktische Anwendungen*. 2. überarb. und aktualisierte Aufl. Bern: Hans Huber, 2011. 278 S. (Wirtschaftspsychologie in Anwendung)
- SCHWEIBENZ, Werner, THISSEN, Frank. *Qualität im Web: benutzerfreundliche Webseiten durch Usability Evaluation*. Berlin: Springer, 2003. 224 S.

Discount Usability-Test

- BARNUM, Carol et al. The „magic number 5“: is it enough for web testing? In: *Extended abstracts on human factors in computing systems (CHI EA '03)* [online]. New York: ACM, 2003. S. 698-699. Verfügbar unter *ACM digital library* (letzter Zugriff: 08.06.2013)
- KRUG, Steve. *Don't make me think!: web usability - das intuitive Web*. 2. Aufl. Heidelberg: mitp, 2006. 203 S.
- NIELSEN, Jakob. Discount usability: 20 years. In: *Nielsen Norman Group* [online]. 14.09.2009. <http://www.nngroup.com/articles/discount-usability-20-years/> (letzter Zugriff: 08.06.2013)
- NIELSEN, Jakob. Why you only need to test with 5 users. In: *Nielsen Norman Group* [online]. 19.03.2000. <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/> (letzter Zugriff: 08.06.2013)
- TECHSMITH. Morae. In: *TechSmith: wir machen Snagit, Camtasia, Jing und andere Software* [online]. 2013. <http://www.techsmith.de/morae.html> (letzter Zugriff: 08.06.2013)

Remote Usability-Test

- ALBERT, Bill, TULLIS, Tom, TEDESCO, Donna. Beyond the usability lab: conducting large-scale online user experience studies. Amsterdam: Morgan Kaufmann, 2010. 310 S.
- BOLT, Nate, TULATHIMUTTE, Tony. *Remote research*. New York: Rosenfeld Media, 2010. 266 S.
- GOUGH, Dabney, PHILLIPS, Holly. Remote online usability testing: why, how, and when to use it. In: *Boxes and arrows* [online]. 09.06.2003. http://www.boxesandarrows.com/view/Remote_online_usability_testing_why_how_and_when_to_use_it (letzter Zugriff: 05.07.2013)
- ETHNIO. Remote usability and UX research tools. In: *Remote usability testing* [online]. 2013. <http://remotereseach.ch/tools> (letzter Zugriff: 17.06.2013)
- LOOP11. Features. In: *Loop11: Remote and online usability testing tool* [online]. 2013. <http://www.loop11.com/features/> (letzter Zugriff: 17.06.2013)
- LOOP11. *Loop11: Remote and online usability testing tool* [online]. 2013. <http://www.loop11.com/> (letzter Zugriff: 08.06.2013)
- SAURO, Jeff. 10 things to know about unmoderated usability testing. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 13.11.2012. <http://www.measuringusability.com/blog/unmoderated-things.php> (letzter Zugriff: 08.06.2013)

Remote vs. Lab testing

- ANDREASEN, Morten Sieker et al. What happen to Remote usability testing?: an emprical study of three methods. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI `07)* [online]. New York: ACM, 2007. S. 1405-1414. Verfügbar unter *ACM digital library* (letzter Zugriff: 08.06.2013)
- BRUUN, Anders et. al. Let your users do the testing: a comparison of three Remote asynchronous usability testing methods. In: *Proceedings of the SIGCHI conference on human factors in computing systems (CHI `09)* [online]. New York: ACM, 2009. S. 1619-1628. Verfügbar unter *ACM digital library* (letzter Zugriff: 08.06.2013)
- KELLY, Diane, GYLLSTROM, Karl. An examination of two delivery modes for interactive search system experiments: remote and laboratory. In: *Proceedings of the SIGCHI conference on human factors in computing systems (CHI `11)* [online]. New York: ACM, 2011. S. 1531-1540. Verfügbar unter *ACM digital library* (letzter Zugriff: 14.07.2013)
- MEIER, Florian, WOLFF, Christian. IA von Websites: asynchrone Remote-Tests und Laborstudien im Vergleich. In: *Mensch und Computer 2012* [online]. München: Oldenbourg Verlag, 2012. S. 133-142. <http://www.oldenbourg-link.com/doi/pdf/10.1524/9783486718782.133> (letzter Zugriff: 08.06.2013)
- SELVARAJ, Prakaash. *Comparative study of synchronous Remote and traditional in-lab usability evaluation methodes* [online]. 2004. 80 S. Master thesis, industrial and systems engineering, Virginia Polytechnic Institute and State University, 2004. http://vtechworks.lib.vt.edu/bitstream/handle/10919/9939/Thesis_Prakaash_Selvaraj.pdf?sequence=2 (letzter Zugriff: 07.06.13)
- TULLIS, Tom et al. *An empirical comparision of lab and Remote usability testing of web sites* [online]. Boston: Fidelity Investments, 2002. 8 S. <http://home.comcast.net/~tomtullis/publications/RemoteVsLab.pdf> (letzter Zugriff: 07.06.13)
- WEST, Ryan, LEHMANN, Katherine R. Automated summative usability studies: an empirical evaluation. *Proceedings of the SIGCHI conference on human factors in computing systems (CHI `06)*. [online]. New York: ACM, 2006. S. 631-639. Verfügbar unter *ACM digital library* (letzter Zugriff: 14.07.2013)

Konzeption und Auswertung der Usability-Tests

- ALBERT, Bill, TULLIS, Tom, TEDESCO, Donna. Beyond the usability lab: conducting large-scale online user experience studies. Amsterdam: Morgan Kaufmann, 2010. 310 S.
- NIELSEN, Jakob, LORANGER, Hoa. *Prioritizing web usability*. California: New Riders, 2006. 406 S.
- SAURO, Jeff. Confidence interval calculator. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 12.03.2012.
<http://www.measuringusability.com/ci-calc.php&summary=1>
(letzter Zugriff: 26.06.2013)
- SAURO, Jeff. Confidence interval calculator for a completion rate. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 01.10.2005.
<http://www.measuringusability.com/wald> (letzter Zugriff: 26.06.2013)
- SAURO, Jeff. Graph and calculator for confidence intervals for task times. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 06.02.2006.
http://www.measuringusability.com/time_intervals (letzter Zugriff: 26.06.2013)
- SAURO, Jeff. If you could only ask one question, use this one. In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 30.10.2012.
<http://www.measuringusability.com/blog/single-question.php>
(letzter Zugriff: 02.03.2010)
- SAURO, Jeff. 10 things to know about the single ease question (Seq) In: *Measuring usability: quantitative usability, statistics and six-sigma* [online]. 30.10.2012.
<http://www.measuringusability.com/blog/seq10.php> (letzter Zugriff: 20.06.2013)
- WIKIA. Repräsentativität. In: *Marktforschungs-Wiki* [online].
<http://marktforschung.wikia.com/wiki/Repräsentativität> (letzter Zugriff: 26.06.2013)

Weitere Quellen

- BUNDESAMT FÜR STATISTIK. Internetnutzung. In: *Statistik Schweiz* [online]. 2013.
http://www.bfs.admin.ch/bfs/portal/de/index/themen/16/04/key/approche_globale.indicator.30106.301.html?open=302,5#5 (letzter Zugriff: 20.06.2013)
- SWISS-LIB. *Swiss-Lib: Schweizerische Diskussionsliste der SpezialistInnen für Information und Dokumentation* [online]. <http://lists.switch.ch/mailman/listinfo/swiss-lib>
(letzter Zugriff: 17.06.2013)

Anhang A: Skript für den Testleiter

Vor dem Eintreffen des Teilnehmers

- Laptop starten
- Browser mit Webseite von Swissbib öffnen
- Morae Recorder öffnen
- Skript für Teilnehmer bereitlegen
- Stift bereitlegen

Empfang des Teilnehmers

- Begrüssung des Teilnehmers und Danksagung
- Getränk anbieten

Vorstellung von Swissbib

Seit 2009 gibt es Swissbib als Betaversion. Swissbib ist ein Metakatalog, der es erlaubt mehrere Bibliothekskataloge gleichzeitig zu durchsuchen. Es ist möglich nach Dokumenten wie Büchern oder CDs zu suchen und zu sehen, wo diese vorhanden sind.

Ziel des Tests

Im Rahmen meiner Bachelorarbeit an der Haute Ecole de Gestion (HEG) Genf vergleiche ich unterschiedliche Methoden zur Evaluation der Benutzerfreundlichkeit am Beispiel des Metakatalogs Swissbib. Ziel der Arbeit ist es, Unterschiede in den Methoden zu identifizieren und Swissbib zu verbessern. Unter anderem führe ich online Benutzerfreundlichkeitstests auf dem Metakatalog Swissbib durch.

Während dem Test sollen nicht Sie, sondern Swissbib getestet werden. Wenn Sie daher auf Probleme stossen, ist das nicht ihr Fehler, sondern liegt daran, dass Swissbib nicht benutzerfreundlich genug gestaltet ist. Ihr Feedback wird mir dabei helfen, Swissbib zu verbessern.

Ablauf

Ich gebe Ihnen nun eine Reihe von Blättern, die sie nacheinander bearbeiten sollen. Fragen können Sie direkt auf dem Papier beantworten. Zum Lösen der Aufgaben steht Ihnen dieser Laptop bereit. Bitte sprechen Sie alle Gedanken laut aus und kommentieren Sie was sie tun. Dies wird mir die Analyse vereinfachen.

Lassen Sie sich Zeit beim Bearbeiten der Aufgaben. Sobald Sie eine Seite gelöst haben oder nicht mehr weiterkommen, blättern Sie um und bearbeiten die nächste Seite. Ich werde im Raum bleiben aber nicht intervenieren. Falls Sie Fragen haben, können Sie mir die jederzeit stellen. Ich werde diese falls möglich beantworten, aber ich werde Ihnen nicht helfen beim Lösen der Aufgaben, damit ich Sie nicht beeinflusse. Falls Sie eine Aufgabe nicht oder nur mit Schwierigkeiten lösen können, ist das nicht schlimm.

Aufnahme und Rechte

Damit eine spätere Analyse möglich ist, wird der Test mit einer Webcam aufgezeichnet und ihre Mausbewegungen auf dem Bildschirm werden gespeichert. Während dem Test zittert die Maus leider ein wenig. Die Teilnahme am Test ist anonym. Alle Daten werden vertraulich behandelt und ausschliesslich zu wissenschaftlichen Zwecken genutzt. Falls Sie sich während dem Test nicht wohlfühlen sollten, können wir jederzeit abbrechen. Ich bitte Sie nun die folgende Einverständniserklärung durchzulesen und falls Sie keine Einwände haben, zu unterschreiben.

Haben Sie noch irgendwelche Fragen? Sonst starten wir jetzt mit dem Test.

Testablauf

- Morae Recorder starten
- Dokumente bereitstellen
- Aufgaben und Fragen bearbeiten
- Morae Recorder stoppen
- Aufnahme speichern

Verabschiedung

- Danksagung
- Gutschein aushändigen

Anhang B: Pretest

Geschlecht

- ☐ - Weiblich
☐ - Männlich

Alter

- ☐ - bis 20
☐ - 21 bis 30
☐ - 31 bis 40
☐ - 41 bis 50
☐ - 51 bis 60
☐ - + 60

Wie erfahren sind Sie in der Nutzung vom Internet?

- | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Gar nicht erfahren | Wenig erfahren | Erfahren | Sehr erfahren | Experte |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Wie erfahren sind Sie in der Nutzung von Bibliothekskatalogen?

- | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Gar nicht erfahren | Wenig erfahren | Erfahren | Sehr erfahren | Experte |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Wie erfahren sind Sie in der Nutzung von Swissbib?

- | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Gar nicht erfahren | Wenig erfahren | Erfahren | Sehr erfahren | Experte |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Anhang C: Aufgaben und Folgefragen

Aufgabe 1

Das Ziel der ersten Aufgabe ist es, die Webseite von Swissbib ein bisschen besser kennenzulernen. Suchen Sie nach einem beliebigen Buch. Nutzen Sie nur Swissbib, um die Frage zu beantworten. Fahren Sie fort, sobald Sie sich einen ersten Eindruck von der Webseite gemacht haben.

Fragen zur Aufgabe 1

Wie war Ihr erster Eindruck von Swissbib? Haben Sie irgendwelche Kommentare zu der Webseite?

Aufgabe 2

Kürzlich haben Sie einen Krimi gelesen, den Sie Ihrem Kollegen gerne weiterempfehlen möchten. Jedoch erinnern Sie sich nur noch an den Titel des Buches, der "Grand Cru" lautete. Suchen Sie den Krimi und merken Sie sich den Autor!

Fragen zur Aufgabe 2

Wie lautet der Autor des Buches "Grand Cru"?

- ☐ - Daniel Black
- ☐ - Martin Walker
- ☐ - Patrick Robinson
- ☐ - Ich konnte die Aufgabe nicht lösen.

Wie schwierig oder einfach fanden Sie diese Aufgabe?

	1	2	3	4	5	6	7	
Sehr schwierig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sehr einfach

Aufgabe 3

Ihre kleine Nichte kommt am Wochenende zu Ihnen nach Chur zu Besuch. Sie möchten sich gerne das Hörspiel "Wickie und die Graumänner" für Ihre Nichte ausleihen. Suchen Sie, ob es dieses Hörspiel in einer Bibliothek in Chur gibt!

Fragen zur Aufgabe 3

Gibt es das Hörbuch "Wickie und die Graumänner" in einer Bibliothek in Chur?

- ☐ - Ja, es gibt das Hörbuch in Chur.
- ☐ - Nein, es gibt das Hörbuch nicht in Chur.
- ☐ - Es gibt überhaupt kein Hörbuch zum Buch "Wickie und die Graumänner".
- ☐ - Ich konnte die Aufgabe nicht lösen.

Wie schwierig oder einfach fanden Sie diese Aufgabe?

	1	2	3	4	5	6	7	
Sehr schwierig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sehr einfach

Aufgabe 4

Versuchen Sie herauszufinden, ob Sie das Hörbuch "Wickie und die Graumänner", welches Sie für Ihre Nichte besorgen möchten, ausleihen können.

Fragen zur Aufgabe 4

Können Sie das Hörbuch ausleihen?

- ☐ - Ja.
- ☐ - Nein, das Hörbuch ist bereits von einer anderen Person ausgeliehen.
- ☐ - Nein, das Hörbuch wird momentan vermisst.
- ☐ - Ich konnte die Aufgabe nicht lösen.

Wie schwierig oder einfach fanden Sie diese Aufgabe?

	1	2	3	4	5	6	7	
Sehr schwierig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sehr einfach

Aufgabe 5

Sie möchten sich gerne mit dem Thema "Bioenergie" kritisch auseinandersetzen. Gibt es eine Möglichkeit herauszufinden, wie viele Dokumente es zu diesem Thema auf Deutsch gibt?

Fragen zur Aufgabe 5

Gibt es eine Möglichkeit, die Anzahl deutscher Dokumente zum Thema "Bioenergie" herauszufinden?

- ☐ - Nein, es gibt keine Möglichkeit.
- ☐ - Ja, es sind ca. 70 Dokumente.
- ☐ - Ja, es sind ca. 300 Dokumente.
- ☐ - Ich konnte die Aufgabe nicht lösen.

Wie schwierig oder einfach fanden Sie diese Aufgabe?

	1	2	3	4	5	6	7	
Sehr schwierig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sehr einfach

Aufgabe 6

Sie haben sich kürzlich ein Konto auf Swissbib eingerichtet und möchten sich nun einloggen. Ihr Login ist "swissbib@safetypost.de" und das Passwort "swissbib".

ACHTUNG: Versuchen Sie zusätzlich herauszufinden, wozu es sinnvoll sein könnte, sich auf Swissbib ein Konto einzurichten (auch wenn Sie sich nicht einloggen konnten).

Fragen zur Aufgabe 6

Wozu könnte es sinnvoll sein, sich ein eigenes Konto auf Swissbib einzurichten?

Wie schwierig oder einfach fanden Sie diese Aufgabe?

	1	2	3	4	5	6	7	
Sehr schwierig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sehr einfach

Anhang D: Posttest-Fragebögen

System Usability Scale

Bitte beurteilen Sie mit Hilfe folgender Aussagen Swissbib:

4. Ich denke, dass ich Swissbib regelmässig benutzen möchte.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

5. Ich fand Swissbib unnötig kompliziert.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

6. Ich fand, Swissbib war einfach zu benutzen.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

7. Ich denke, ich würde die Unterstützung eines Experten benötigen, um Swissbib nutzen zu können.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

8. Ich fand, die verschiedenen Funktionen von Swissbib spielten gut zusammen.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

9. Ich fand, dass es auf Swissbib zu viele Unstimmigkeiten gab.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

10. Ich kann mir vorstellen, dass die meisten Leute die Benutzung von Swissbib sehr schnell lernen.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

11. Ich fand die Nutzung von Swissbib sehr umständlich.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

12. Ich fühlte mich sicher im Umgang mit Swissbib.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

13. Ich musste einiges lernen, bevor ich mit Swissbib zurechtkam.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

Zweiter Posttest-Fragebogen

Bitte beurteilen Sie mit Hilfe folgender Aussagen die Testmethode:

14. Im Allgemeinen fühlte ich mich während des Tests sehr wohl.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

15. Die Aufgabenstellungen waren gut erklärt.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

16. Der Aufwand für die Teilnahme am Test war gross.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

17. Die Aufgaben waren realitätsnah.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

18. Ich fühlte mich unter Druck, die Aufgaben richtig zu lösen.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

19. Ich kann mir vorstellen, auch in Zukunft an solchen Tests teilzunehmen.

	1	2	3	4	5	
Trifft nicht zu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Trifft voll zu

Haben Sie Kommentare zu Swissbib oder zur Testmethode?

Anhang E: Einverständniserklärung - In-person Test

Diese Einverständniserklärung wurde im Rahmen einer Bachelorarbeit an der Haute Ecole de Gestion (HEG) Genf verfasst. Ziel der Arbeit ist es, unterschiedliche Methoden zur Evaluation der Benutzerfreundlichkeit am Beispiel des Metakatalogs Swissbib zu vergleichen, um Unterschiede in den Methoden zu identifizieren und Swissbib zu verbessern.

Für die Arbeit werden Benutzerfreundlichkeitstests durchgeführt. Dabei werden Gesicht, Ton und Bildschirm aufgezeichnet. Die Teilnahme am Test ist anonym. Alle gewonnenen Daten werden vertraulich behandelt und ausschliesslich zu wissenschaftlichen Zwecken genutzt und Ton- oder Bildaufnahmen werden nicht veröffentlicht. Der gesamte Test oder einzelne Aufgaben dürfen jederzeit abgebrochen werden.

☐ - Ich bin nicht damit einverstanden, dass ich mit einer Kamera aufgenommen werde.

Ich wurde über meine Rechte und Pflichten aufgeklärt und erkläre mich hiermit mit diesen einverstanden.

Vorname Name: _____

Datum: _____

Unterschrift: _____

Anhang F: Rekrutierungs E-Mail - Remote Test

Haben Sie schon einmal an einem Online-Benutzerfreundlichkeits-Test teilgenommen? Dies ist Ihre Möglichkeit!

Im Rahmen meiner Bachelorarbeit an der Haute Ecole de Gestion (HEG) Genf vergleiche ich Methoden, die es erlauben, die Benutzerfreundlichkeit von Webseiten zu evaluieren.

Primäres Ziel der Arbeit ist es, Unterschiede in den Evaluationsmethoden zu identifizieren. Dazu werde ich einen traditionellen Usability-Test mit einem Online-Usability-Test vergleichen. Die Tests werden auf dem Metakatalog **Swissbib** durchgeführt, welcher es erlaubt, mehrere Bibliothekskataloge gleichzeitig zu durchsuchen.

Für die Online-Tests suche ich nach Freiwilligen. Sie können den Test **überall und wann immer** Sie wollen durchführen. Es dauert ca. 15 bis 20 Minuten. Jeder Teilnehmer mit Wohnsitz in der Schweiz erhält ein **kleines Dankeschön**.

Habe ich Ihr Interesse geweckt?

Zögern Sie nicht, hier geht's direkt zu Usability-Test:

<http://www.loop11.com/usability-test/18477/introduction/>

Alle Daten werden vertraulich behandelt und ausschliesslich zu wissenschaftlichen Zwecken genutzt.

Falls Sie gerne mehr über den Test erfahren möchten oder an den Resultaten der Studie interessiert sind, können Sie mich jeder Zeit unter folgender E-Mail-Adresse kontaktieren:

[REDACTED]

Eveline Schmidt

Studentin zweisprachiger Studiengang Information und Dokumentation

HEG Genf

Anhang G: Ausschnitte aus Loop11 - Remote Test

Begrüssung



Benutzerfreundlichkeits-Test von Swissbib


Herzlichen Dank, dass Sie beschlossen haben, an diesem Test teilzunehmen! Ihr Feedback wird mir dabei helfen, meine Bachelorarbeit zu verfassen. Jeder Teilnehmer mit Wohnsitz in der Schweiz erhält ein kleines DANKESCHÖN!

Die Teilnahme am Test ist anonym. Alle Daten werden vertraulich behandelt und ausschliesslich zu wissenschaftlichen Zwecken genutzt.

Viel Spass beim Durchführen des Tests wünscht Ihnen
Eveline Schmidt

Evaluation starten 

Standard-Erklärung




Benutzerfreundlichkeits-Test von Swissbib


Im Rahmen dieser Studie werden Sie aufgefordert, verschiedene Aufgaben auszuführen.


- Die jeweilige Aufgabe wird am oberen Seitenrand angezeigt, wie im Folgenden dargestellt.


20%
abgeschlossen

Example: Using the website below, find the contact details of the organisation

 Aufgabe abbrechen


Aufgabe abgeschlossen 
- Um die einzelnen Aufgaben zu erfüllen, müssen Sie innerhalb der Website zu der Seite navigieren, die die Informationen für diese Aufgabe enthält bzw. zu der Seite, deren Inhalt Ihnen am besten geeignet scheint, um die Aufgabe zu erfüllen.
- Klicken Sie in der Menüleiste am oberen Seitenrand auf ‚Aufgabe abgeschlossen‘, wenn die Seite mit den Informationen für die Aufgabe angezeigt wird.

Aufgabe abgeschlossen 
- Wenn Sie die entsprechende Seite nicht finden bzw. Schwierigkeiten haben, die Aufgabe zu lösen, dann wählen Sie ‚Abbrechen‘.

 Aufgabe abbrechen

Keine Angst: wir testen nicht Sie, sondern die Website.

Klicken Sie auf WEITER, wenn Sie diese Anweisungen gelesen haben, um mit der ersten Aufgabe zu beginnen.

WEITER 

Beispiel: Pretest-Fragebogen

0%
abgeschlossen

Geschlecht

☐ Weiblich

☐ Männlich

WEITER →


Beispiel: Aufgabe 2

25%
abgeschlossen

Kürzlich haben Sie einen Krimi gelesen, den Sie Ihrem Kollegen gerne weiterempfehlen möchten. Jedoch erinnern Sie sich nur noch an den Titel des Buches, der "Grand Cru" lautete. Suchen Sie den Krimi und merken Sie sich den Autor!

Abbrechen Aufgabe abgeschlossen →

DE EN FR IT



Home My swissbib Über swissbib Feedback ? Hilfe Anmelden

Suchen nach alle Bibliotheken Suchen → Erweiterte Suche

29%
abgeschlossen

Wie lautet der Autor des Buches "Grand Cru"?

☐ Daniel Black

☐ Martin Walker

☐ Patrick Robinson

☐ Ich konnte die Aufgabe nicht lösen.

WEITER →

33%
abgeschlossen

Wie schwierig oder einfach fanden Sie diese Aufgabe?

☐ Sehr schwierig ☐ ☐ ☐ ☐ ☒ ☐ Sehr einfach

WEITER →

Beispiel: Posttest-Fragebogen

88%
abgeschlossen

Beurteilen Sie mit Hilfe folgender Aussagen den Katalog Swissbib:

Ich denke, dass ich Swissbib regelmässig benutzen möchte.

☐

1

☐

2

☐

3

☐

4

☐

5

Trifft nicht zu

Trifft voll zu

Ich fand Swissbib unnötig kompliziert.

☐

1

☐

2

☐

3

☐

4

☐

5

Trifft nicht zu

Trifft voll zu

Ich fand, Swissbib war einfach zu benutzen.

☐

1

☐

2

☐

3

☐

4

☐

5

Trifft nicht zu

Trifft voll zu

Verabschiedung

Loop¹¹


Benutzerfreundlichkeits-Test von Swissbib

Herzlichen Dank! Sie haben den Test nun beendet.

Damit Sie Ihr Dankeschön erhalten, schreiben Sie mir doch bitte eine Mail an: . Gerne informiere ich Sie auch über die Resultate der Studie, falls Sie interessiert sind.

Ich danke Ihnen recht herzlich für die Teilnahme am Test!

Eveline Schmidt

Studie abgeschlossen 

Anhang H: Usability-Probleme

Nr.	Problem	Schweregrad	In-person	Remote
3	Übernahme der Autovervollständigung	  	✓	✗
4	Unrelevante Vorschläge, Null-Treffermengen	  	✗	✓
2	Dublettenbereinigung	  	✓	✓
22	Lesbarkeit im Menü „Hilfe“	  	✓	✓
24	Anmeldeverfahren	  	✓	✓
5	Verarbeitung von Umlauten	  	✗	✓
6	Vorschläge bei Null-Treffermengen zu klein	  	✗	✓
8	Sortierung nach Relevanz	  	✓	✓
9	Unauffällige Fassettennavigation	  	✓	✓
12	Unklares Merksymbol	  	✓	✗
15	„Zum Bestand“ & „Informationen zum Bestand“	  	✓	✓
16	Bezeichnung „Zum Bestand“	  	✓	✗
18	Anzeige der Verfügbarkeit (Internet Explorer)	  	✓	✗
20	Organisation der Hilfe-Themen	  	✗	✓
21	Navigation innerhalb der „Hilfe“	  	✓	✗
26	Festlegen der Favorit-Bibliotheken	  	✓	✓
27	Bezeichnung „Temporäre Merkliste“	  	✓	✗
1	Grösse der Menüleiste	  	✓	✗
2	Autovervollständigung verschwindet	  	✓	✗
10	Öffnen des Dokuments mit Bild	  	✓	✗
11	Bibliothek identifizieren	  	✓	✗
13	Unklarer Link zur erweiterten Suche	  	✓	✗
14	Bezeichnung des Formats „Sprache“	  	✓	✓
17	Informationen über die Bibliothek	  	✗	✓
19	Anzeige des Marc-Formats (Internet Explorer)	  	✓	✗
23	Unterschied „My swissbib“ und „Anmelden“	  	✓	✓
25	Zugriff auf das Benutzerkonto	  	✓	✗